Piero Mella

The Magic Ring

Systems Thinking Approach to Control Systems

Springer

Contemporary Systems Thinking

Piero Mella

The Magic Ring

Systems Thinking Approach
to Control Systems

Springer

Piero Mella
Department of Economics and Management
University of Pavia
Pavia, Italy

Preface

*Thinking is easy, acting is difficult, and to put one's thoughts
into action is the most difficult thing in the world
(Johann Wolfgang Von Goethe).*

THE ORIGINS—This book represents the completion and evolution of my book *Systems Thinking. Intelligence in Action* (Springer 2012), which presented the general view of the "world" considered as a "system of systems" of interconnected variables. That book presented the theory, language, and basic rules of *Systems Thinking,* while this book develops the theory, language, and rules behind the functioning of *Control Systems*, which were presented in a single short chapter. Though this book is the logical development of the previous book, it is a self-contained work compared to *Systems Thinking* since it does not suppose any prior knowledge of Systems Thinking—whose basic concepts are presented at the beginning—but gradually introduces a true Discipline of Control Systems that is accessible to a large group of nonspecialists. The choice of topics and examples—in this endless field of knowledge—derives from the need to make the Discipline of Control accessible to "everyone."

THE ASSUMPTIONS—The idea of developing a Discipline of Control that stands on its own with regard to Systems Thinking is based on the hypothesis—whose validity I have tried to demonstrate throughout the book—that among all types of systems Control Systems occupies an absolutely preeminent position. Even if we are not accustomed to "seeing them," "recognizing them," or "designing them," they are everywhere inside and around us. Only their ubiquitous presence makes possible our world, life, society, and existence, constructing an ordered and livable world, erecting barriers against disorder, and directing the irregular dynamics of the world toward states of homeostatic equilibrium in all its vital and social environments. It is for this reason I have tried to make clear the need to gradually accustom ourselves to thinking not only in terms of systems but above all in terms of Control Systems: "*from Systems Thinking to Control Thinking*"; this is my epistemological proposal.

THE OBJECTIVES—This book presents a gradual path toward "educating" the reader in understanding how Control Systems truly operate and in recognizing, simulating,

and improving them in all fields of activity. Starting from the hypothesis that knowledge of Control Systems is not only a technical fact but also represents a discipline—that is, "A discipline is a developmental path for acquiring certain skills or competencies. (. . .) To practice a discipline is to be a lifelong learner. You "never arrive"; you spend your life mastering disciplines." (Senge, 1992, p. 12)—I have set the objective of making Control Systems a topic that is, in a certain sense, simple and attractive by turning to the effective symbolism typical of Systems Thinking models and avoiding too technical and formal a treatment of the subject. Thus the reader should know that this is not an engineering, physics, biology, or economics text, nor a mathematics one either. Technical or mathematical tools are not used to construct Control Systems; instead, I use the highly simple and universal logic behind the notion itself of *control process* and the simple and universal action of the Control Systems that produce this process.

THE TITLE—WHY "RING?"—Control Systems (in their most simple form) are particular apparatuses (logical or technical) that cause a variable Y_t to attain an objective Y^* by gradually reducing the error (gap, distance) $E_t = Y^* - Y_t$ through a series of "adjustments" of Y_t produced by acting on a *control variable*, or *lever*, X_t, so that the values of Y_t approach Y^*; the iterations of this process continue and, if the system is well designed, E_t becomes zero, even in the presence of external disturbances, D_t, which cannot be controlled by the system. Using the logical language of Systems Thinking we can easily see that a control system is a typical balancing loop that connects the three variables [X_t, Y_t, and E_t]; this loop is represented by a typical circular graphic model that in fact resembles a *Ring*. We can also imagine that each iteration to reduce E_t is like a *rotation* of the *Ring*.

THE TITLE—WHY "MAGIC?"—It is probably difficult to realize that you can read these sentences *only* because at least five fundamental automatic Control Systems, which exist in our bodies to produce vision, are acting simultaneously; that is, five *Rings* that rotate at an incredible speed. I will briefly mention these, and without any particular order (for simplicity's sake I shall omit the time reference in the variables).

The first *Ring* is the focus Control System, which works through the crystalline lens which, by varying its shape using the tensor muscles (X), varies the focal distance in order to eliminate the gap between the actual focus (Y) and that needed to read the letter (Y^*), even when the distance between the text and the reader's eye (D) varies. The second *Ring* is the control system which, by changing the diameter of the pupil (X), varies the quantity of light (Y) that strikes the retina (cones and rods) to maintain (Y^*) constant even with variations in the ambient light (D). The third *Ring* is the control system that, by using the motor eye muscles (X), produces the macro and micro movements in the eyeball (Y) in order to read the words in the line (Y^*), if necessary aided by the rotation of the head (controlled by an additional *Ring*). The fourth *Ring* is the control system that produces tears (X) to keep the cornea wet (Y) in an optimal manner (Y^*), even when there is a change in evaporation due to heat at various times from various sources (D). The fifth *Ring* (the final one I shall mention) is perhaps the most important: the Control System

that allows us to *recognize* signs written on a page as *letters*, and the latter as *words* given rhythm by punctuation marks, which leads to the *recognition* of sentences.

Assuming that these five *Rings alone* "rotate" *only* 20 times per second, then in order to read this paragraph in, let's say 60 s, we need 6,000 total iterations ($5 \times 20 \times 60$), that is, 6,000 control actions of which we are normally unaware: the number, in fact, required to read second by second (obviously ignoring whether or not the reader understands what he has read). Does this number of iterations seem too small? Consider that obviously these five *Rings* function continuously all day long our entire lives: let us assume on average 14 h for 80 years. During this period the five Control Systems produce more than (145×10^9) iterations. Our five *Rings* rotate at an incredible speed only to guarantee vision, one of the five senses. Just think about the others!

Does it seem like a small number now?

The most complex "machine" on earth is probably the human brain, formed on average by ($100 \times 10^9 = 10^{11}$) neurons (in addition to an even larger number of auxiliary cells). The complexity of this machine derives from the fact that the neurons are connected by means of various types of synapses, whose overall number is estimated to be 10^{14}. By using the simple model of the neural networks we can imagine that the synapses have a particular "weight" which determines the strength of the signal received by each transmitter neuron which is then sent to the receptor neurons, and that each neuron in turn has a specific threshold value below which the neuron does not "fire" the signals received, causing the connection with the receptor neurons down the line not to take place. It is immediately clear that for each neuron there must correspond "at least one" Control System for identifying the threshold value and that for each synapse "at least one" Control System for sending the signals with their specific weights. At this moment our brain is working thanks to the action of ($10^{11} + 10^{14}$) *Rings* rotating thousands of times per second, and for the entire human population [$(10^{11} + 10^{14}) \times 6 \times 10^9$] neural *Rings* are operating (in the brain). We may not realize it, but this is "magic".

Does this seem too small a number?

The average number of cells in the body of a man weighing 70 kg is estimated to be 10^{14}. Each cell, as an autopoietic and thus homeostatic system, can exist thanks to the action of a large number of Control Systems that govern the cell cycle, instant by instant, and to the other Control Systems that regulate cellular interactions. Estimating that there are *only* 100 Control Systems for each cell, at this moment your body can remain alive thanks to the action of 10^{16} invisible *Rings* that rotate uninterruptedly for your entire life. At this moment the human population exists thanks to the action of [$6 \times 10^9 \times 10^{16}$] cellular Control Systems in addition to [$(10^{11} + 10^{14}) \times 6 \times 10^9$] neural Control Systems, trillions of trillions of invisible *Rings* that rotate incessantly, untiringly, second after second, year after year. I am sure few people have considered this.

Do you still think this is not significant?

Without exaggerating and becoming tiresome, I would point out that the cellular control characterizes all cells of all living beings. Ants alone are estimated to number from 10^{15} to 10^{16}. How many insects are there? How many birds? How

many mammals? How many fish and mollusks? It is estimated that there are more than (400×10^6) units of phytoplankton in a cubic meter of water. The oceans and seas are made up of about (1.34×10^9) km^3 of water, the equivalent of (1.34×10^{18}) cubic meters. Overall there are $(400 \times 1.34 \times 10^{24})$ individual plankton, single- or multicelled, which at this moment are kept alive by $(k \times 400 \times 1.34 \times 10^{24})$ invisible *Rings* that rotate in whirling fashion second after second. It is too difficult, useless even, to estimate the number of vegetal units—plants, shrubs, blades of grass, mushrooms, lichen, mould, etc.—and the number of cells of each individual. However, there is no need for such an estimate to grasp the impressive number of *Rings* necessary to keep the planet's flora alive.

And what about the billions of man-made Control Systems? And the trillions of Control Systems that regulate relationships between individuals, populations, and species? And the endless number of Control Systems that operate in our atmosphere, seas, and in all physical phenomena?

There is a lot still to discover about how Control Systems function as motors for learning and knowledge; I will deal with this in Chap. 9 on.

If there is something magical in life, society, and organization, in orders and evolution, this something is represented by the "magic" action of the "rings" which, through their ubiquitous presence and variety of forms, account for individuality, maintain the order of things, restore equilibrium and regulate interactions.

Why "magic *Ring*"? Independently of its nature, function, and physical structure, each Control System always has the same structure of a *Ring*. A "magic *Ring*," in fact.

The CONTENT—The book is divided into ten Chapters.

Chapter 1 seeks to review the basic language of Systems Thinking and the models it allows us to create. This language and these models will be the instruments for presenting the logic of Control Systems in a manner that is understandable to everyone. It is suggested that the study of Control Systems be considered as a true discipline, according to Peter Senge's conception; a true Sixth Discipline.

Chapter 2 introduces the control process, presenting the *theoretical structure* of three simple Control Systems we all can observe: those that permit us to regulate the sound volume (radio, TV, iPod, stereo, computer, etc.), the water temperature (shower, radiator, boiler, etc.), and the room temperature (air conditioner, refrigerator, etc.). Despite the simplicity of the three kinds of systems we can gain some fundamental knowledge from them about the basic structure of a Control System by learning the basic concepts: *manager, governor, action rate, reaction rate, reaction time, delays,* and *disturbances.* After presenting the general theoretical model of this system, we examine several new concepts: (1) the *chain of control*, which represents the technical structure of the process; (2) the various types of *delays* in control, extending the considerations already begun in Chap. 2; (3) *interferences* among the Control Systems which can give rise to disturbance dynamics.

Chapter 3 proposes a *general typology* of Control Systems with examples taken from observations of reality, distinguishing among: (a) systems of *attainment* and *recognition*; (b) *continuous-* or *discrete-lever systems*; (c) *fixed* or *variable objective* systems; (d) *steering* and *halt* Control Systems; (e) systems of *collision,*

anti-collision and *queuing* systems; (f) *tendential* and *combinatory* systems; to name but a few.

Chapter 4 broadens the view of Control Systems by introducing two important generalizations: (1) *multilever* Control Systems, with levers that are independent or dependent of each other; (2) *multi-objective* systems, with independent or interdependent objectives. In multilever systems it is fundamental to understand the importance of the control *strategy*; that is, of the specification of the priorities for the activation of the different levers used to eliminate the deviation from the objective. In multi-objective systems, on the other hand, it is fundamental to understand the importance of the control *policy*, which decides on the order of priorities for intervening on the various objectives. To further generalize, we shall examine the notion of *impulse systems*, which play a fundamental role in life.

Chapter 5 outlines the guidelines for recognizing, observing or designing Control Systems and presents the problems that arise regarding their logical realization. It will introduce the fundamental distinction between symptomatic and structural control.

Chapters 6–9 undertake a "mental journey" through various "environments," increasingly broader in scope, suggesting to the reader how to recognize therein Control Systems that, by their ubiquitous presence, make the world possible in all its manifestations.

Chapter 6 considers daily individual actions—the domestic and civic environment (the most familiar ones)—in order to arrive at Control Systems that maintain the psychic–physical system.

Chapter 7 guides the reader through the physical, biological, and social environments. Particular attention is given to Control Systems that operate in Societies and in *combinatory systems*, a special class of systems which has not yet been adequately examined.

Chapter 8 considers the environment of organizations, trying to demonstrate how organizations are themselves Control Systems by presenting two basic models: Stafford Beer's Viable System Model and my own Model of the Organization as an Efficient System of Transformation (MOEST). The chapter also illustrates the logic behind the many Control Systems that operate in the network of processes in organizations and which are necessary to maintain autopoiesis and increase the efficiency of the entire organization.

Chapter 9 deals with three specific topics: (1) the *cognition process*, understood as the result of the action of Control Systems that recognize differences and order them based on Gregory Bateson's notion of "mind"; (2) the *signification process* which allows man to translate cognitive content into signs and to form languages; (3) the *scientific process* and the explanatory power of models.

Chapter 10, the final chapter, deals with two topics: (1) some ideas about a Discipline of Control Systems are proposed; (2) the human aspects of control. A reflection on the content and limits of this book are presented as FAQs in the final pages of the Chapter.

THE STYLE AND METHOD—In order to facilitate the understanding, acquisition and practice of the discipline, Control Systems have been presented in a logical way by

using the symbols of Systems Thinking to produce easy-to-construct models. Thus I have intentionally avoided expressing Control Systems in a typically mathematical form, which would have used differential equations or differences, preferring instead to use simulation examples with "everyday resources," in particular Excel files accompanied (in a few cases) by Powersim, to demonstrate to the reader how relatively easy it is to do the calculations to achieve the control. Moreover, I have chosen particularly simple and clear examples by presenting very common Control Systems we all can observe and experiment with in order to draw lessons in general theory from them.

THE READERSHIP—The book is "for everyone." There is no prerequisite required to read and understand it, in particular math and computer knowledge. The use of the simplest tools of Excel helps in creating models for basic simulations of simple systems in order to improve understanding. However, I have felt it useful to also present several simple models in Powersim. A number of examples are aimed at systems found in organizations and companies; thus the text can aid in the professional growth of managers, consultants, and corporate employees in general. Teachers, professionals, and educated people in general can also find sources for reflection and occupational tools from the text. A word of caution for the reader: *in order to master the Control Discipline it is necessary to read this book "slowly" and refer the standard models presented to one's own world and experiences.*

Pavia, Italy   Piero Mella

Contents

List of Figures

List of Tables

Part I
Discovering the "Ring"

Chapter 1
The Language of Systems Thinking
for Control Systems

Systems thinking plays a dominant role in a wide range of fields from industrial enterprise and armaments to esoteric topics of pure science. Innumerable publications, conferences, symposia and courses are devoted to it. Professions and jobs have appeared in recent years which, unknown a short while ago, go under names such as systems design, systems analysis, systems engineering and others. They are the very nucleus of a new technology and technocracy; their practitioners are the "new utopians" of our time [...] who—in contrast to the classic breed whose ideas remained between the covers of books-are at work creating a New World, brave or otherwise (Ludwig von Bertalanffy 1968, p. 3).

The great shock of twentieth-century science has been that systems cannot be understood by analysis. The properties of the parts are not intrinsic properties but can be understood only within the context of the larger whole. Thus the relationship between the parts and the whole has been reversed. In the systems approach the properties of the parts can be understood only from the organization of the whole. Accordingly, systems thinking concentrates not on basic building blocks, but on basic principles of organization. Systems thinking is "contextual," which is the opposite of analytical thinking. Analysis means taking something apart in order to understand it; systems thinking means putting it into the context of a larger whole (Fritjof Capra 1996b, 29–30).

We are about to begin a wonderful trip in the world of control systems, which are all around us. We are not used to seeing them, but their presence is indispensable for the existence of the world and of life. We need to gradually understand how they really operate in order to recognize, simulate, and improve them. The reader should know that this is not a mathematics, engineering, physics, biology, economics, or sociology book. It does not treat mathematical or technical tools used to produce or even simulate factual control systems.

I have chosen to present the logic of control systems using the simple but powerful and clear language of systems thinking for interpreting the world as a system of dynamic systems of interconnected and interacting variables. The first chapter thus serves to recall the principles of systems thinking and present the rules of the formal language used to represent dynamic systems, the language we will also use to construct models of the control systems that will be developed in subsequent chapters. The pervasiveness of the logic, notions, and techniques that

can be applied to the multiform variety of control systems allows us to develop a true "art" of control—a "discipline"—which I shall call the *Sixth Discipline, since it represents the evolution of the way* systems thinking (the "fifth discipline") *observes the world*.

1.1 The Sixth Discipline: The Discipline of Control Systems

In his important work, *The Fifth Discipline: The Art and Practice of the Learning Organization* (Senge 1990, 2006; Senge et al. 1994), which is aimed mainly at managers, professionals, and educated people in general, Peter Senge seeks to spread systems thinking by applying a simple and clear methodology and symbology that everyone can easily understand and which permits us to build qualitative models of the world understood as the interpenetration—to different extents—of systems of interacting variables (Mella 2012) of all types of structure and complexity that often form holarchies or holonic networks (Mella 2009).

Senge indicates to managers the "superiority" of systems thinking not in order to propose a new form of thinking but to emphasize the need for organizations and managers to think systematically. Precisely for this reason, he considers systems thinking as a discipline—the fifth discipline—everyone should follow in order to get accustomed to succinct thinking, the only kind that can understand the properties of a whole without, however, forgetting about those of the parts.

In fact, the history of systems thinking, as a form of holistic thinking, is very old, as Fritjof Capra shows in his important work *The Web of Life*:

> The ideas set forth by organismic biologists during the first half of the century helped to give birth to a new way of thinking—"systems thinking"—in terms of connectedness, relationships, context. According to the systems view, the essential properties of an organism, or living system, are properties of the whole, which none of the parts have. They arise from the interactions and relationships among the parts (Capra 1996b, p. 30).

In this sense Senge views systems thinking as an efficient method for abandoning reductionist thinking, which blocks us from seeing the emergent properties of systems—above all, dynamic systems—and from learning to think in a succinct manner as well.

Rather than concentrate on mathematical and technical formalism—which can be found, if necessary, in specialist texts on systems theory (Bertalanffy 1968; Boulding 1956, Klir 1969; Casti 1985; Skyttner 2005), systems science (László 1983, Sandquist 1985; Klir 1991), systems analysis and design (Wasson 2006; Gopal 2002), control theory and cybernetics (Ashby 1957; Wiener 1961; Johnson et al. 1963; Leigh 2004), and other fields—Senge wished to codify a way of thinking about dynamics and complexity—that is, a systems thinking—that presents an effective approach for observing and modeling reality, forcing the observer to search for the circular connections, the loops among the interrelated variables of which reality is composed. More than a technique, systems thinking is thus a mental disposition (Anderson and Johnson 1997, p. 20), an attitude, a language, and a paradigm.

> Systems Thinking [is] a way of thinking about, and a language for describing and understanding, the forces and interrelationships that shape the behavior of Systems. This discipline helps us see how to change systems more effectively, and to act more in tune with the larger processes of the natural and economic world (Senge et al. 1994, p. 6).

> Systems Thinking is a Paradigm and a Learning Method. The first conditions the second. The second supports the first. The two parts form a synergistic whole (Richmond 1994, online).

> "Systems Thinking" has long been a cornerstone in our work on organizational learning, but the term often seems more daunting (it can easily sound like an intellectual task reserved for Ph.D.'s) than helpful. In fact, systems thinking is not about fighting complexity with more complexity. It simply means stepping back and seeing patterns that are, when seen clearly, intuitive and easy to grasp (Senge et al. 2008, p. 23).

It is for this reason that it must be gradually learned and systematically applied; it must become a true discipline that every company man, as well as each of us, must master through experience.

> Systems thinking is a discipline for seeing wholes, recognizing patterns and interrelationships, and learning how to structure those interrelationships in more effective, efficient ways (Senge and Lannon-Kim 1991, p. 24).

Other terms can also be used to indicate this new way of thinking. Barry Richmond, who was one of the most renowned experts in this discipline (he was the founder, in 1984, of the High-Performance System), stated:

> Systems Thinking, a Systems Approach, Systems Dynamics, Systems Theory and just plain ol' "Systems" are but a few of the many names commonly attached to a field of endeavor that most people have heard something about, many seem to feel a need for, and few really understand. [. . .] As I prefer the term "Systems Thinking," I'll use it throughout as the single descriptor for this field of endeavor (Richmond 1991, p. 1).

Systems thinking has been proposed as the "fifth discipline" precisely because, together with four other disciplines, it is presented as a means for building learning organizations, which are organizations that develop a continual collective learning by putting all their members in a position to learn together while supplying them with the instruments for such collective learning.

> Systems thinking is the fifth discipline. [. . .] It is the discipline that integrates the [other four] disciplines, fusing them into a coherent body of theory and practice. [. . .] Without a systemic orientation, there is no motivation to look at how the disciplines interrelate. By enhancing each of the other disciplines, it continually reminds us that the whole can exceed the sum of its parts (Senge 2006, p. 12).

Among the most notable results of systems thinking, the most important in my view is its ability to allow us to understand, explain, and simulate the *modus operandi* of control systems, which are the basic elements of biological and social life; of populations, societies, and organizations; and even of most physical phenomena. It is not difficult to see that without the action of control systems, which, though by nature extremely varied, can all be represented by the same models, the "world" itself—understood as a system of increasingly more extensive systems— would probably not exist, if the variables it is made up of were not in some way limited by constraints and objectives that keep the system "what it is" within a time horizon that preserves the relations among the variables.

Why, then, do we need a new book exclusively dedicated to control systems? The reason I propose this book is to demonstrate, using the logical and symbolic apparatus of systems thinking, that all control systems, no matter the context they operate in, can be represented by the same "logical models" and obey a relatively simple theory whose understanding does not require a particular knowledge of math, physics, or engineering. Thus, this is not an engineering, physics, biology, or economics book, nor for that matter a math book; neither technical tools (Sontag 1998; de Carvalho 1993; Goodwin et al. 2001; Leigh 2004) nor mathematical equations (Vanecek and Celikovsky 1996; Imboden and Pfenninger 2013) necessary to construct sophisticated control systems are considered. My sole aim is to present the highly simple logic underlying the concept of control, which is understood as the gradual approximation of the movement of a variable toward desired values (objectives), made possible by the action of one or more additional variables—the control levers—whose values are recalculated, based on some specific time scale, in order to gradually counter any movement away from the objective.

Thus, I also propose a more ambitious objective: to arouse in the reader curiosity about control systems and help him or her recognize these in the various environments they operate in, thereby providing a guide to the construction of models useful for understanding and simulating their behavior. I will not follow particular techniques but merely help the reader get accustomed to the logic of control and to work toward mastering them, slowly guiding him or her toward increasingly vaster systems. In other words, I will introduce him or her to what I feel could become a true discipline of control systems—or, more simply, "control discipline"—in line with Peter Senge's understanding of the term "discipline."

> By "discipline", I do not mean an "enforced order" or "means of punishment," but a body of theory and technique that must be studied and mastered to be put into practice. A discipline is a developmental path for acquiring certain skills or competencies.
>
> A discipline is a developmental path for acquiring certain skills or competencies. As with any discipline, from playing the piano to electrical engineering, some people have an innate "gift," but anyone can develop proficiency through practice.
>
> To practice a discipline is to be a lifelong learner. You "never arrive"; you spend your life mastering disciplines (Senge 2006, p. 10).

Following Senge's fifth discipline, I hope not to be too pretentious in calling the discipline of control systems the sixth discipline, that is: the discipline of the control of the individual, the collectivity, organizations, and the ecosystem. In other words, the discipline of our world's present and future. In fact, in Chap. 19 (entitled appropriately enough "A Sixth Discipline?") of the first edition of *The Fifth Discipline* Senge, in the specific context of organizations and companies, writes:

> The five disciplines now converging appear to comprise a critical mass. They make building learning organizations a systematic undertaking, rather than a matter of happenstance. But there will be other innovations in the future. [. . .] perhaps one or two developments emerging in seemingly unlikely places, will lead to a wholly new discipline that we cannot even grasp today. [. . .] Likewise, the immediate task is to master the possibilities presented by the present learning discipline, to establish a foundation for the future (Senge 1990, p. 363).

It truly is difficult to think we can lay the foundations for the future without understanding the logic of the control systems that allow us to maintain intact over time our organism, society, and the organizations we are a part of, which, through their processes, generate stability and order as well as development and innovation. I believe that the need for a new culture of control is also recognized by Peter Senge in his recent book *The Necessary Revolution*, which applies systems thinking to the socioeconomic system in order to explore the paths of change in industrial civilizations, a change held to be indispensable for guaranteeing the *sustainability* of our social and economic system in a context where common exhaustible resources are becoming increasingly scarcer.

> A sustainable world, too, will only be possible by thinking differently. With nature and not machines as their inspiration, today's innovators are showing how to create a different future by learning how to see the larger systems of which they are a part and to foster collaboration across every imaginable boundary. These core capabilities—seeing systems, collaborating across boundaries, and creating versus problem solving—form the underpinnings, and ultimately the tools and methods, for this shift in thinking (Senge et al. 2008, p. 11).

Since change is only possible by applying the logic of control systems, *change management*, which decides how to regulate the system in order to achieve the new objectives, must necessarily identify the structure and state of the actual socioeconomic system along with the levers to modify the objectives through the levers of cultural change, which lead to a "*working together to create a sustainable world*" (for more on the processes of change management, see Sect. 7.11).

As Fritjof Capra states, evolution, innovation, and learning are also typical of control processes as regards the vision provided by cybernetics as well as the more comprehensive view of self-organizing systems:

> For Ashby [in his Cybernetic view] all possible structural changes take place within a given "variety pool" of structures, and the survival chances of the system depend on the richness, or "requisite variety," of that pool. There is no creativity, no development, no evolution. The later models, [Self-organization view] by contrast, include the creation of novel structures and modes of behavior in the processes of development, learning, and evolution (Capra 1996b, p. 85).

How can we consciously guide the present toward the future, proactively face the unavoidable consequences of the processes under way—global warming, extinction, population growth, spread of epidemics, arms escalation, space exploration, biological and nanotechnological research, inflation, poverty, etc.—and consciously react to new events that we cannot foresee today (cataclysms, harmful innovations, etc.) if we are not able to understand the logic, power, or networks of control systems in action and, above all, to design and build effective control systems? How can we innovate, create, and manage new forms of organization, new machines, and new processes without also conceiving new control systems? If we think about it carefully, even the purely evolutionary process "produces" nothing other than a continuous innovation in control systems by "producing" new sensors (antennae, eyes, electrical discharge sensors, etc.), new regulatory organs (the nervous system and brain in particular), and new effectors of movement and of the recognition of forms (Sect. 2.13). Where are the control systems? They are

within us and everywhere around us. We must be able to identify them by zooming out to catch the extreme variety, richness, and importance of the macro control systems as well as zooming in to catch the infallible effectiveness of the micro control systems, which are so essential to life.

Though different, all control systems can be referred to a few model types that are particularly useful in all fields of research. We need to get used to observing them and learn to recognize and master them by trying to understand their *modus operandi*. We will realize that we are formed by control systems, and surrounded by control systems, that we can exist and live thanks only to the control systems that regulate our environment and entire ecosystem. When we apply the sixth discipline, our life will no longer seem the same, and perhaps—with the same certainty as John von Neumann (the first citation at the beginning of Chap. 2)—we shall succeed in controlling—after having understood and identified the most suitable control systems—even the most unstable phenomena that are leading us toward the threshold of chaos.

Before examining the logic of control and control systems it is useful to recall the basic principles of the fifth discipline, systems thinking (Sects. 1.2–1.4), and touch on the other four disciplines (Sect. 1.6.7). At this point the importance and pervasiveness of control systems and the relevance of the sixth discipline will be even clearer.

1.2 The Fifth Discipline: The Five Basic Rules of Systems Thinking

I am convinced that intelligence depends on the ability to construct coherent and sensible models (Mella 2012) in order to acknowledge and understand the world. I believe that:

> *Intelligent* persons are those who understand (and comprehend) quickly and effectively; who are not content to "look at the world with their eyes" (objects, facts, phenomena and processes) but who are able "to see the world with their minds" by constructing models to "understand" how the world is (description), how it functions (simulation), and how we can act as part of it (decision and planning), even without having the need, or possibility, of "looking at everything" (Mella 2012, p. 3).

> If the capacity to *see* and not simply look at [the world] depends on the ability to construct models to understand, explain and simulate the world, then the most useful and effective models to strengthen our intelligence are the systems ones based on the logic of Systems Thinking (*ibidem*, p. 6),

since these models allow us to represent complexity, dynamics, and change (the relevance of models for our knowledge and intelligence is shown in Chap. 9, Sect. 9.14). The cognitive efficiency of models derived from systems thinking owes to their ease of construction; they require only perspicacity and acumen; rely on elementary techniques; are understandable even for non-experts; can easily be communicated, debated, and improved; and without too much difficulty turned into quantitative simulation models based on system dynamics logic. Of course we

should not be content only with systems thinking models, but for those with little time or resources to construct more sophisticated (though less immediate) models the following proverb always applies: "*Beati monoculi in terra caecorum*" [blessed are those who see with one eye in the land of the blind]; that is, in a dynamic and complex world blessed are those who, knowing how to construct systems thinking models, have at least one eye in a land of blind people (Mella 2012: Sect. 4.13). However, before examining the technique for constructing models it is useful to recall what systems thinking entails and what its logical and theoretical bases are. Even if Peter Senge has not explicitly stated them, five fundamental rules are at the basis of systems thinking in my opinion, as I have tried to demonstrate in my book *Systems Thinking* (2012, Chap. 1).

FIRST RULE: If we want to understand the world we must be able to "see the trees and the forest"; we must develop the capacity to "zoom" from the whole to the parts, from systems to components, and vice versa.

This rule, which is at the basis of systems thinking, can be translated as follows: Reality is *permeated with systems*, increasingly vaster in scope, which form a global structure that creates a global process that cannot be *understood only* from the *outside* or the *inside*. This rule, whose mastery requires constant practice, can be translated: If we want to broaden our intelligence we *must develop the capacity to "zoom" from* parts *to* the whole *and from* entities *to* components. In this sense we can say (Mella 2009) that this FIRST RULE of systems thinking "operationalizes" the *holonic view* introduced in 1967 by Arthur Koestler in his book *The Ghost in the Machine*.

> Parts and wholes in an absolute sense do not exist in the domain of life [...] The organism is to be regarded as a multi-levelled hierarchy of semi-autonomous sub-wholes, branching into sub-wholes of a lower order, and so on. Sub-wholes on any level of the hierarchy are referred to as holons [...] The concept of holon is intended to reconcile the atomistic and holistic approaches (Koestler 1967, p. 341).

To understand the dynamics of the world we must always remember that each "object" to which our observation is directed is at the same time a whole—composed of smaller parts—and part of a larger whole: a holon! Ken Wilber, another contemporary proponent of the holonic view (Wilber 2000), is even more explicit, stating: "*The world is not composed of atoms or symbols or cells or concepts. It is composed of holons*" (Wilber 2001, p. 21). I shall return to the holonic view of the world in more detail in Sect. 3.8.

SECOND RULE: We must not limit our observation to that which appears constant but "search for what varies"; the variables and their variations are what interest the systems thinker.

This rule, equally as important as the preceding one, imposes greater attention and a more intense discipline, since systems thinking tells us to shift from a "world of objects" (people, processes, concepts, etc.)—whether trees or forests—to a "world of variables" that connote those objects. By definition a variable X may assume different values (measurements, states, quantity, etc.): $x_0, x_1, x_2, x_3, \ldots$ which we can assign through a succession of measurements. Let $x(t_0)$, $x(t_1)$, $x(t_2)$, $x(t_3)$, etc.

represent the values of X measured at the end of regular intervals, $[t_1, t_2, t_3, t_4, \ldots]$, all within a defined period of reference T. In this sense X is a (discrete) time variable and its measured values, arranged along a time axis, form the dynamics (motion, trajectory, evolution, trend, etc.) of X with respect to T. Systems thinking is interested above all in the variations they reflect: $\Delta x(t_1) = x(t_1) - x(t_0)$; $\Delta x(t_2) = x(t_2) - x(t_1)$; $\Delta x(t_3) = x(t_3) - x(t_2)$; etc. This seems easier than it really is. One needs sensitivity and experience to select the truly significant variables. Systems thinking explores and models systems of interconnected and interacting variables which produce the complexity we observe in our world.

> Senge sorts two types of complexity from this—detail complexity and dynamic complexity. Detail complexity arises where there are many variables, which are difficult, if not impossible, to hold in the mind at once and appreciate as a whole. Dynamic complexity arises where effects over time of interrelatedness are subtle and the results of actions are not obvious; or where short term and long term effects are significantly different; [...] (Flood 1999, p. 13).

However, we must not limit ourselves to explicitly stating the variables we consider useful but must be able to measure the "variations" they undergo over time. For systems thinking, each object must be observed as a "vector of variables," and the dynamics of the objects are perceived only if we are able to observe and measure the dynamics of the variables they represent. Systems thinking also requires us to *select* the relevant variables and *restrict* their number so as to consider only those most relevant for the construction of models. In general, a few attempts are sufficient to narrow down the most interesting variables.

THIRD RULE: *If we truly wish to understand reality and change we must make an effort "to understand the cause of the variations in the variables we observe" by forming chains of causal relationships among the connected variables while identifying and specifying:*

1. The *processes* that "produce" the dynamics in the variables and the *machines* (or *systemic structures*) that "produce" those *processes*
2. The *variables* that "carry out" those processes (causes or inputs) and those that "derive" from the processes (effects or outputs)

For simplicity's sake we could even call the input and output variables "causes" (causal variables) and "effects" (caused variables), respectively. However, we must always remember that the processes—to the extent they are considered *black boxes*—always play the role of producer of the *effects*, given the *causes* and the more or less broad set of "initial" and "boundary or surrounding" conditions.

This third rule is perhaps the most important. It assumes the constant exercise of the two preceding ones while also completing them, forming together a very potent logical system: "*if you want to understand the world you mustn't be content with only observing variables and their variations, but you must search for the 'cause' of the variations in the variables you observe*." This seems quite obvious, but the fact that normally we do not observe, or else fail to consider, that the variables change their values due to some process carried out by some "machine" makes it equally important to practice applying this rule, which has three corollaries:

(a) Every observable variable is a *cause* or an *effect* of other variables.
(b) Cause and effect variables are linked to form causal chains (A causes B, B causes C, . . .) where each variable (A, B, C, etc.) is both the effect of variables up the line and a cause for those down the line.
(c) It is not possible to understand the world's dynamics without identifying and studying the causal chains that make up the world.

FOURTH RULE: It is not enough to search for the causes of the variations we observe by only searching for the causal chains among variables; we must also "link together the variables in order to specify the loops *among all the variations."*

A loop is formed when two variables, *X* and *Y*, are linked in a *dual direction*: *X* causes *Y* and *Y* causes *X* (naturally we must take into account the time sequence of the observations) so that the concept of *cause* and *effect*—which holds between two (or more) linked variables forming a causal chain—loses its significance when we consider those variables linked together by one or more *loops* (Richardson 1991). For this reason systems thinking states: if we really want to "understand" the world and its changes, it is not enough to reason in terms of chains of causes and effects between variables; we must instead get used to *circular* thinking. In other words, we must be aware that every variable is not only the cause of variations in variables connected down the line in the causal chain but that the causal chain can become a *loop*, since the last variable in the chain can be the cause of the first variable, which becomes "the effect of its effects," thereby creating a circular connection. In conclusion, this rule obliges us to make an effort to link together the variables until we obtain a loop among their variations.

In other words, we must move from the causal chains to the systemic *interconnections* and from the *linear variations* to the *systemic interactions* among the variables of interest. In brief, we must see the world in terms of circular processes, or feedback loops, abandoning the "linear thinking" ("laundry list thinking") that only considers chains of causes and effects and getting accustomed to "circular thinking" (loops and causal loop diagrams, CLDs), identifying the loops that connect the variables.

> [. . .] If you took the time to record your thoughts, I'll bet they took the form of a [. . .] "laundry list". I like to refer to the mental modelling process that produces such lists as laundry list thinking. I believe it to be the dominant thinking paradigm in most of the Western world today. [. . .] Notice that the implicit assumptions in the laundry list thinking process are that (1) each factor contributes as a cause to the effect, i.e., causality runs one way; (2) each factor acts independently; (3) the weighting factor of each is fixed; and (4) the way in which each factor works to cause the effect is left implicit (represented only by the sign of the coefficients, i.e., this factor has a positive or a negative influence). The systems thinking paradigm offers alternatives to each of these assumptions. First, according to this paradigm, each of the causes is linked in a circular process to both the effect and to each of the other causes. Systems thinkers refer to such circular processes as feedback loops (Richmond 1993, p. 117).

A set of variables linked in various ways by loops so that the behavior of any variable depends on that of all the others forms a *dynamic system*, according to systems thinking. Systems thinking defines "system" as a unitary set of

interconnected variables—capable of producing emerging macro-dynamics that do not coincide with any of the micro-dynamics of the individual variables or their partial subsystems—whose logical structure it investigates and represents (Mella 2012, p. 21). The chains and loops that bind together the variables form the *logical structure of the system*. Systems thinking states that to "describe, model, and control" the world it is enough to understand the logical structure of the systems of variables it is composed of, leaving to engineers, physicists, biologists, doctors, economists, sociologists, psychologists, and other specialized scientists the task of investigating the operating structure of those systems.

At this point we can now finally derive a general law of systems thinking: "*in order to explain the dynamics of a variable, do not search for its immediate cause but rather define and investigate the logical structure of the dynamic system it is a part of.*" But how large must a dynamic system be? How many variables do we have to connect? To answer these questions systems thinking proposes this final rule.

FIFTH RULE: *When we observe the world we must "always specify the boundaries of the system we wish to investigate."*

In fact, the FIRST RULE of systems thinking obliges us to zoom inside a system— thereby identifying increasingly smaller subsystems—as well as outside a system to identify ever larger supersystems.

Are we thus destined (or "condemned") to having a holistic view without limits? Do we have to zoom *ad infinitum*? Certainly not! Systems thinking is the art of "seeing" the world, and in order for what we see to have a true meaning it must depend on our cognitive interests. We cannot have a forest without limits. For this reason the fifth rule of systems thinking obliges us to always specify the boundaries of the system being investigated, since our capacity to understand a system is diminished when it becomes too large. In fact there are two boundaries: an *external* boundary which delimits the system when we zoom from the parts to the whole and an *internal* one when we zoom from the whole to the parts. It is not easy to identify or set the system's boundaries; fortunately, the more we apply ourselves in the discipline of systems thinking, the more the solution to this problem becomes easy, almost spontaneous.

1.3   The Construction of Models Based on Systems Thinking: The *Rings*

As I have observed, systems thinking, or the fifth discipline, has the great merit of making immediately and easily understandable the *causal* interrelationships among the variables that make up a *dynamic* system, understood as a *pattern* of *interconnected* and *interacting* variables of whatever type, size, and structure. In order to make the causal relations among the variables in a dynamic system immediately and easily understandable, systems thinking constructs models by following five operational rules, which can be summed up as follows (Mella 2012).

1. Once the cause and effect links between two variables (for example, X and Y) have been determined, the link is represented by simply using an arrow (of whatever shape, thickness, or color) that unequivocally correlates their variations. The cause variable (input) is written at the tail of the arrow and the effect variable (output) at the head. To indicate that X causes Y, systems thinking writes:

$$
\begin{array}{cc}
\textbf{cause} & \textbf{effect} \\
X \xrightarrow{\text{causal relationship}} & Y \\
x(t) & y(t) \\
\Delta x & \Delta y \\
\mathrm{d}x & \mathrm{d}y
\end{array}
$$
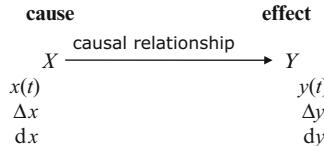
The direction of the arrow indicates the direction of the causal link, so that

$$
X \xrightarrow{\text{causal relationship}} Y
$$

is different than

$$
Y \xrightarrow{\text{causal relationship}} X
$$
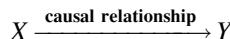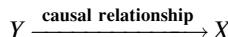
2. After having specified the direction of the cause and effect link we need to understand the direction of variation of this link; two variables have the *same direction* of variation (**s**) if increases or decreases $(+/-)$ in the former result in corresponding increases or decreases $(+/-)$ in the latter. They have the *opposite direction* (**o**) if increases or decreases in the former $(+/-)$ result in corresponding decreases or increases $(-/+)$ in the latter. The direction of variation is indicated by placing (**s**) or (**o**) next to the arrow, as shown in the following models:

$$
\underset{+/-}{X} \xrightarrow{\ \textbf{s}\ } \underset{+/-}{Y} \qquad \underset{+/-}{Y} \xrightarrow{\ \textbf{o}\ } \underset{-/+}{X}
$$

3. Identify the causal chains, the significant variables these are composed of, and the direction of variation of the entire chain. Each causal chain has a direction of variation "**s**" or "**o**" that depends on the directions of variation of the causal links they are made up of. If there is an odd number of variations "**o**" among the variables in the chain, then the entire chain has a direction of variation "**o**"; an increase (decrease) in the *first* variable $(+/-)$ produces a decrease (increase) in the *last* one $(-/+)$; if there are no "**o**" variations, or an even number of them, then the causal chain has direction "**s**."

4. In many cases two variables—let us suppose X and Y—can be linked by cause and effect relations that act contemporaneously in two opposite directions. We must recognize and represent the *two-way* circular cause and effect links, that is, those between X and Y and, at the same time, between Y and X. This possible dual relation is represented by two opposite arrows—from X to Y and

Y to *X*—which form the simplest model of a dynamic system, called the CLD or causal diagram. Loops can be *basic*, when there are only two variables, or *compound*, when more than two variables are joined in a circular link. The loops are distinguished by the effect produced on the variables they are composed of, as a result of the direction of variation of the variables themselves. There are only two basic types of loop:

(a) *Reinforcing* loops [**R**], which produce a *reciprocal increase* or *reduction*—in successive repetitions of the system's cycle—in the values of the two variables having an identical *direction of variation*: "**s** and **s**" or "**o** and **o**"; a basic reinforcing loop can be represented by the following model:
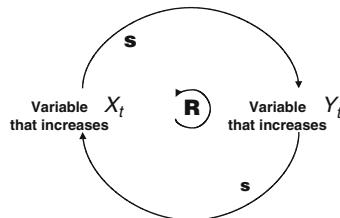
(b) *Balancing* loops [**B**], which *maintain relatively stable* the values of the connected variables, which are connected by a different *direction of variation*: "**s** and **o**" or "**o** and **s**"; a basic balancing loop can be represented as follows:
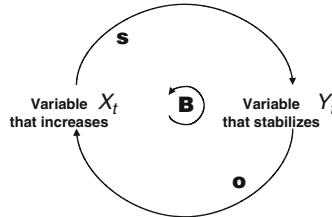
I propose referring to a *basic* or a *compound loop* with the more evocative term *"Ring."*

5. Join together the various reinforcing and balancing loops in order to represent any type of reality formed by the interconnected variables included within the system's boundary. *A system of loops in which all variables are linked by arrows, without there being an initial and final variable, is defined as a general CLD.*

By applying these simple operational rules, systems thinking states that a dynamic world, no matter how large and complex, can be effectively described and modeled through various combinations of loops [**R**] and [**B**], even by inserting in each of the loops a large number of variables linked by causal relations (zooming in) or by adding any number of loops through some variables (zooming out). Each general CLD represents a model of a *dynamic system* in which the dynamics of each variable depends on the dynamics of the others, in addition to variables outside the
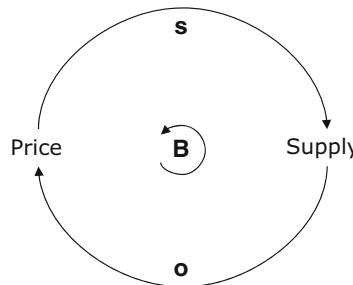
Fig. 1.1 Basic *Ring* connecting price and supply
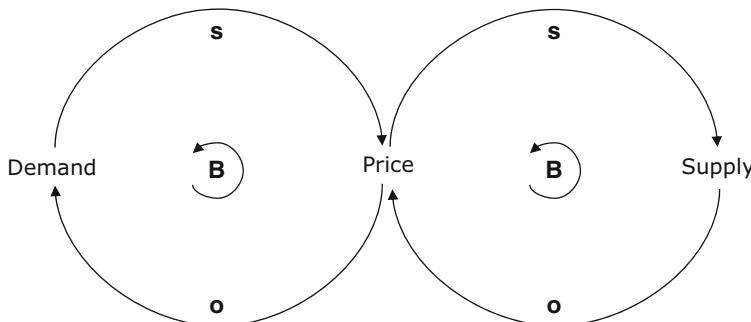
Fig. 1.2 The law of the market under perfect competition

system, which, in the CLD, are represented with arrows whose origin is not linked to other arrows.

To clarify operating rules A–E above, it is useful to present some simple models of dynamic systems. The price of a good has a direction of variation contrary to that of the volumes of supply (downward arrow, "**o**"), but the price variations produce a variation of same direction in the production volumes (upward arrow, "**s**"), as shown in Fig. 1.1. The causal model that links price and supply shows how, in successive *cycles of the system*, an equilibrium [**B**] is produced among the variables; if the price falls too much then supply falls; if supply declines then the price increases again.

The model in Fig. 1.2 shows that it is possible to join two balancing loops to form a system that describes the well-known law of the market under perfect competition; price represents the key variable that produces an equilibrium between supply and demand.

Observing the two joint *Rings* in Fig. 1.2 we can easily see the dynamics of the variables: if the price falls too much then supply will also fall and demand rise, so that even the price increases again, inverting the direction of variation in demand and supply. The CLD in Fig. 1.3 instead represents a reinforcing loop [**R**], since each of the variables causes an increase in the second one (this example was presented in Senge); in successive repetitions of the process, the variation in both variables is accentuated.
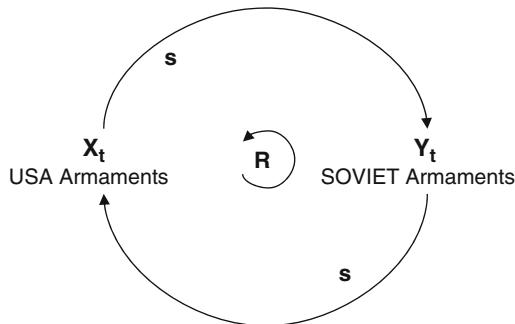
Fig. 1.3 Arms escalation
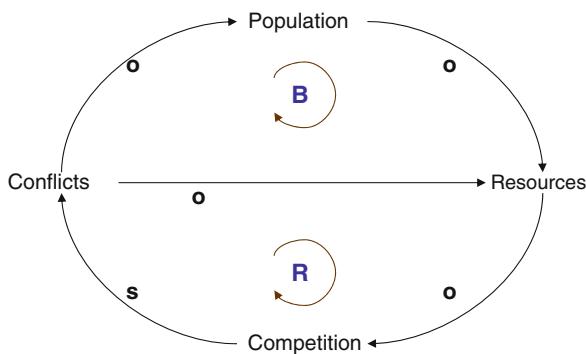
Fig. 1.4 Struggle for life
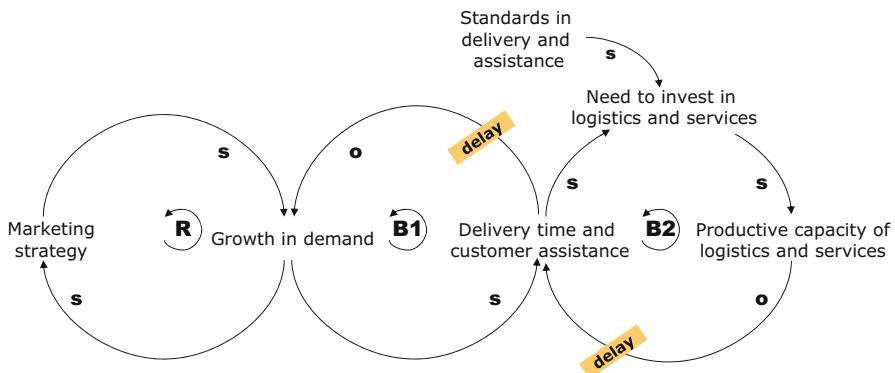
Fig. 1.5 Marketing strategy

The *Ring* in Fig. 1.4 represents a composite, though simplified, model of the system that regulates the struggle for life among populations, groups, and individuals.

The CLD in Fig. 1.5 connects the main variables that drive the marketing strategy.
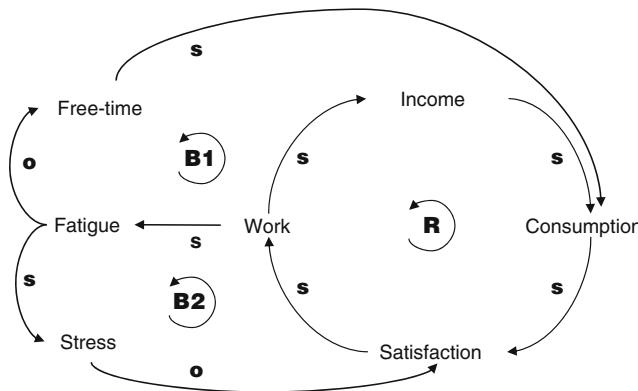
Fig. 1.6 Job satisfaction
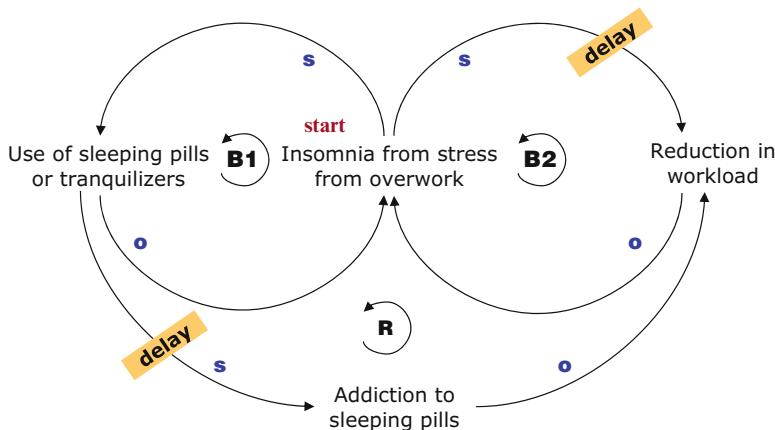
Fig. 1.7 Burnout from stress

Figure 1.6 shows the structure that regulates our job and wealth aspirations.

Figure 1.7 illustrates the system that describes burnout from stress; the search for short-term symptomatic solutions for stress can aggravate the problem in the long term rather than solve it.

Now that we have summarized the five rules of systems thinking and represented the structure of some simple systems, we can understand the approach of the fifth discipline in observing the interconnected systems of variables. As we know, Ludwig von Bertalanffy provided the following general definition of system:

A system can be defined as a complex of interacting elements. Interaction means that the elements, p, stand in relations, R, so that the behavior of an element p in R is different from its behavior in another relation R′. If the behaviors of R and R′ are not different, there is no interaction and the elements behave independently with respect to the relations R and R′ (von Bertalanffy 1968, p. 55).

This definition implies that each system has its own emergent characteristics compared to those of its constituent parts. Thus the definition of dynamic system as a dynamic unity composed of interconnecting and interacting variables is justified. Naturally, a dynamic system derives from a system of processors that produce the dynamics in the input (causes) and output (effects) variables.

Along with dynamic systems there are *structural* systems and *ordinal* systems. *Structural systems* are operational systems made up of elements (usually processors) connected by a map of defined and stable relations forming the invariant organization of the system, according to Maturana and Varela's interpretation (1980) so clearly described by Stafford Beer in the Preface to their work.

> The relations between components that define a composite unity (system) as a composite unity of a particular kind, constitute its organization. In this definition of organization the components are viewed only in relation to their participation in the constitution of the unity (whole) that they integrate. This is why nothing is said in it about the properties that the components of a particular unity may have other than those required by the realization of the organization of the unity (Beer 1992, p. XIX).

For example, systems of production facilities in a factory, organs that are part of the organization of corporations, battalions, companies, and platoons in a regiment are all structural. Clearly every dynamic system has its own structural system, concisely indicated as the "technical structure" of the system. *Ordinal systems* are units made up of non-dynamic elements (objects, concepts, words, etc.) arranged in an ordered map: the Eiffel Tower, made up of the ordered arrangement of structural components; this book, composed of the ordered arrangement of chapters, sections, sentences, and words; and the "afterlife" system described in Dante's Divine Comedy, with its ordered arrangement of the nine circles of hell, the seven terraces of purgatory, and the nine celestial spheres of Paradise. Every dynamic system is characterized by its own *structural system* that defines the system's "logical structure" and is represented by the CLDs.

Systems can be observed from an *external point of view* (forest) following a typical synthetic, teleological, or instrumental approach, according to which these systems, viewed as a whole, in many cases denote an object-oriented macro behavior, in that they serve a particular aim, or from an *internal point of view* (trees), following a structuralist, analytical approach, according to which it detects the macro behavior of the unitary system as the emergent effect of the micro behaviors.

The rules of systems thinking and the simple examples presented clearly show that the *fifth discipline has mainly an analytical, internal approach*; the system derives from the interconnected action of the variables that form the structure. However, it is equally apparent that, after having identified the "trees," we must be able to observe the "forest" by identifying the boundary of the system observed as a whole, according to the synthetic approach, trying to analyze the emerging effects of the overall system. Many systems in systems thinking have no objective and no clear-cut instrumentality; but they are nevertheless wholes that possess their own logical structure and macro behavior and that are investigated following a holistic and holonic approach.

1.4   From Systems Thinking to System Dynamics: A Simulation of a Dynamic System

From the brief considerations in the preceding *sections* it should be clear that systems thinking, or the fifth discipline, considers solely *dynamic systems* that have loops among the variables, so that the latter, being mutually linked, form "a whole" that presents its own dynamics, *new* and *emergent*, with respect to the dynamics of the component variables.

Systems thinking models are typically suited to qualitative research into the systems represented; they offer an understanding of their structure and allow us to examine the direction of the dynamics of the variables. These same models can, however, be translated into simulation models that allow us to numerically calculate and represent in tables and graphs the succession of values of the variables produced by the action and structure of the system under study. Expressed quantitatively, systems thinking adopts the simulation procedures of system dynamics, a discipline and technique that unquestionably goes back to Jay Forrester and his fundamental book *Industrial Dynamics* (Forrester 1961). In recent works Forrester defines systems dynamics as follows:

> By "systems thinking" I mean the very popular process of talking about systems, agreeing that systems are important, and believing that intuition will lead to effective decisions. Without a foundation of systems principles, simulation, and an experimental approach, systems thinking runs the risk of being superficial, ineffective, and prone to arriving at counterproductive conclusions...

> "System dynamics" is a professional field that deals with the complexity of systems. System dynamics is the necessary foundation underlying effective thinking about systems. System dynamics deals with how things change through time, which covers most of what most people find important. System dynamics involves interpreting real life systems into computer simulation models that allow one to see how the structure and decision-making policies in a system create its behavior (Forrester 1999, p. 1).

> System dynamics combines the theory, methods, and philosophy needed to analyze the behavior of systems not only in management, but also in environmental change, politics, economic behavior, medicine, engineering, and other fields. System dynamics provides a common foundation that can be applied wherever we want to understand and influence how things change through time. The system dynamics process starts from a problem to be solved—a situation that needs to be better understood, or an undesirable behavior that is to be corrected or avoided. The first step is to tap the wealth of information that people possess in their heads. [...] System dynamics uses concepts drawn from the field of feedback control to organize available information into computer simulation models (Forrester 1991, p. 5).

Let us consider the *basic loops*, or *Rings*, connecting two variables—for example, X, cause, and Y, effect—which change their values over a time interval $T = [t = 0, t = 1, t = 2, \ldots]$ divided into intervals of equal length, thereby producing reciprocal differences—ΔX_t and ΔY_t—through two opposing and linked *processes*. There are different ways to specify the *functions* needed to simulate the joint dynamics of X and Y (in both directions), since it is possible to variously define the equations that link the variations of X to those of Y.

First, the time sequence of the variations must be accurately determined; for example, the values ΔX_t could produce a variation ΔY_t at the same instant "t" (which is the normal assumption) or a variation ΔY_{t+1} quantified at instant "$t+1$," a variation ΔY_{t+2}, or even other variants; in the opposite direction, a variation ΔY_t could quantify a variation ΔX_{t+1} (as is normally the case) or present other variants.

Second, systems thinking reminds us that the processes that produce the variations in X and Y are carried out by a set of interconnected operative "apparatuses," or "machines," of a particular type and nature, each of which can be viewed as a black box whose *modus operandi* can even be ignored, unless we wish to zoom inside them. To simulate the dynamics of X_t and Y_t we must know how each "operative machine" transforms the variations in the cause variable, X, and the variations in the other effect variable, Y. Therefore, we must assume that each "operative machine" generates the dynamics of the variables by means of a *variation function* that makes the values of the effect variable, Y_t, depend on those of the cause variable, X_t. This implies that for each simulation we must define the two *variation functions*:

1. The function for variations in Y due to variations in X, which we will call $g(Y/X)$, so that an input ΔX_t *always* produces, for example, the output:

$$\Delta Y_{t+1} = [g(Y/X)\,\Delta X_t] + DY_t$$

2. The function for variations in X due to variations in Y, which we will call $h(X/Y)$, so that an input ΔY_t *always* produces, for example, the output:

$$\Delta X_{t+1} = [h(X/Y)\,\Delta Y_t] + DX_t$$

The variables DY_t and DX_t indicate "disturbances" to the two variables from the environment outside the loop. We can represent the loop connecting X and Y in a dual direction, for example using the two joint equations, which show a process length of one period:

$$\begin{cases} \Delta Y_t = \left[g(Y/X)\Delta X_t\right] + DY_t \\ \Delta X_{t+1} = \left[h(X/Y)\Delta Y_t\right] + DX_{t+1} \end{cases} \tag{1.1}$$

If we assume that the structure of the "machine" does not vary over time and does not depend on the values of X or Y, then we can also assume, by simplifying, that the activity of the "machines" that connect X_t and Y_t is represented by an "action parameter," which we indicate by "$g = $ const" for X and "$h = $ const" for Y, so that we can represent the loop (at least in the simplest cases) using two joint equations:

$$\begin{cases} \Delta Y_t = [g\Delta X_t] + DY_t \\ \Delta X_{t+1} = [h\Delta Y_t] + DX_{t+1} \end{cases} \tag{1.2}$$
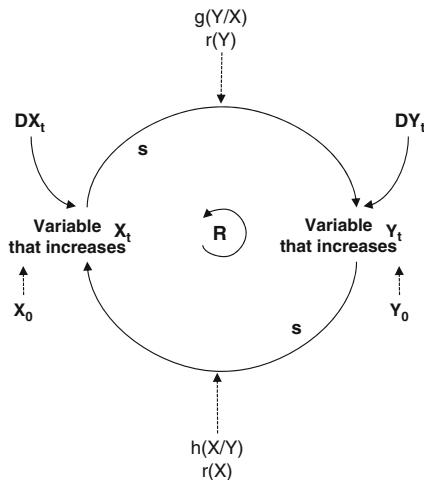
Fig. 1.8 Structure of a reinforcing *Ring* for simulations
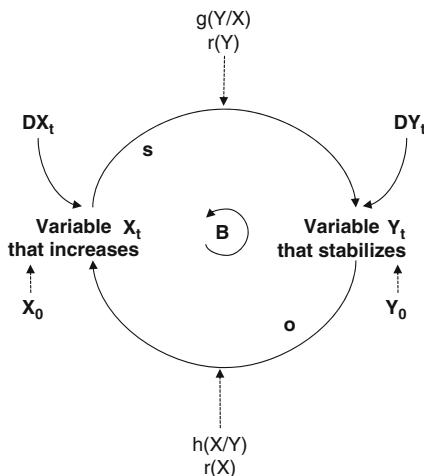
Fig. 1.9 Structure of a balancing *Ring* for simulations

I propose to call "*g*" the "*rate of action*" and "*h*" the "*rate of reaction*." However, to carry out the simulations we must introduce into the model, in addition to the action parameters "*g*" and "*h*" (or their more complex formulations), the *initial values* X_0 and Y_0 of the two variables, along with any limits of variation admissible for them, or *variation constraints*. Figures 1.8 and 1.9 illustrate how the reinforcing "rings" and balancing "rings" can be structured to carry out simple simulations.

To make clearer the dynamics of the reinforcing loops we can assume that, in a simplified way, X and Y of the loop model [**R**] in Fig. 1.8 are linked by the following equations:

$$\begin{cases} X_{t+1} = X_t + \dfrac{Y_t \times h(Y/X)}{r(Y/X)} + DX_t \\[3mm] Y_{t+1} = Y_t + \dfrac{X_{t+1} \times g(X/Y)}{r(X/Y)} + DY_t \\[3mm] t = 1, 2, \ldots \end{cases} \qquad (1.3)$$

where

$$\begin{cases} \Delta Y_{t+1} = Y_{t+1} - Y_t \\ \Delta X_{t+1} = X_{t+1} - X_t \end{cases} \qquad (1.4)$$

In addition to the functions describing the variation of the first variable as a function of the second, equations (1.3) introduce the variable "r" to indicate the *reaction time*, which represents the number of time scales over which the adjustment of variations in Y due to variations in X (of variations in X due to variations in Y) takes place. Referring to Y, a value $r = 2$ means that the variation in Y is halved in each interval since it is distributed over two intervals (I will take this up in more detail below). If "g" and "h" represent the *action* and *reaction rates* (which do not depend on the value of the causal variables X and Y), $r = 1$, and there are no disturbances, then the system normally has a linear dynamics; if, instead, "g" and "h" are themselves functions of X and Y and/or of time and/or $r = 2, 3, \ldots$ then the dynamics of the system will very likely be nonlinear.

A loop [**B**] can be represented in many ways, either by acting on the signs of the action rates "g" and "h" or on the signs of the causal links; in any event, we can construct the following recursive functions:

$$\begin{cases} X_{t+1} = X_t - \dfrac{Y_t \times h(Y/X)}{r(Y/X)} + DX_t \\[3mm] Y_{t+1} = Y_t + \dfrac{X_{t+1} \times g(X/Y)}{r(X/Y)} + DY_t \\[3mm] t = 1, 2, \ldots \end{cases} \qquad (1.5)$$

As Fig. 1.9 clearly shows, a process in the balancing *loops* leads to a variation in the direction "**s**" of the Y_t, but this causes a variation in the opposite direction, "**o**," in X_t; in this case, an impulse (an increase or a decrease in X_t) is "balanced" by the opposite variation in Y_t.

The two variables *balance* each other; in other words, a positive impulse ΔX_t causes an incremental variation ΔY_t. The new value of Y_{t+1} produces a decreasing

variation in X_{t+1}, so that if the processes occur over several cycles, the two variables will tend to reciprocally compensate their variations and the balancing loop will produce cyclical dynamics in the variables.

There are no other elementary *Rings* since a loop where both variables vary in direction "**o**" would still be a reinforcing loop entirely similar to that in the model in Fig. 1.8.

1.5   Two Fundamental Laws of Systems Thinking

Given that we can represent a dynamic world of interconnected variables, building CLDs makes clear that *systems thinking* represents a very powerful and efficient way of thinking that enhances our intelligence. Systems thinking must be considered, however, not only as a technique but primarily as a *discipline* for efficient and effective thinking, in that it proposes:

– To train us to observe reality as composed of dynamic systems
– To provide us powerful models of description and simulation
– To improve our ability to acquire knowledge, that is, to learn
– To develop our intelligence

Systems thinking, precisely because it is a tool for developing our intelligence, must be learned gradually through practice and constant improvement (Checkland 1999). It is a discipline that requires the systems thinker to have a deep knowledge and to constantly apply its rules as well as to have the willingness to continually improve. This objective can be achieved by developing the necessary competencies in order to:

– Perceive and recognize the circularity of phenomena
– See how systems really function
– Intuit the effects of actions over time
– Feel responsible for the system's performance
– Predict the future
– Control dynamic processes

It is important to clarify which systems systems thinking examines and what types of models can thereby be obtained. As we have clearly specified, due to its intrinsic logic systems thinking considers *dynamic systems* of any kind in any field, building models of a world of incessant movement in continual transformation and evolution. This discipline considers not only dynamic but also *repetitive* systems, which are able to repeat their processes over time, as well as *recursive* systems, capable of interacting with themselves, in the sense that their output, entirely or in part, becomes their own inputs. Even if we are not used to observing them, repetitive and recursive systems are all around us; they are the typical essence of nature. Life itself is recursive in its typical process of birth, reproduction, and death, which is destined to repeat itself again and again.

Systems thinking is particularly sensitive to *systems with memory*; it forces us to consider the connections between variables, always zooming between high-level variables, which accumulate variation over time, and more detailed (state) variables, which cause variations over time; it forces us to observe the dynamics of recursive processes and not only individual pairs of values to consider the loops and not only the pure causal connections (von Foerster 2003, p. 143).

With its core of general principles and concepts, the systems thinking discipline makes it possible (though not easy, since it is a discipline) "to see" an interconnected, dynamic, and recursive world, that is, to identify the structure that determines the trends in phenomena and produces complexity, nonlinearity, and memory in systems that operate in all fields of human action, biological, economic, and social. A good model of the world under examination must always foresee balancing factors for the reinforcing loops and reinforcing factors for the balancing loops.

The systems thinking models are certainly not the only ones capable of increasing our knowledge of the world, but in my view their cognitive effectiveness owes to their ease of construction and communication (Nonaka 1994). The only skills they require are perspicacity and insight; they use elementary techniques, are understandable even to non-experts, and can easily be communicated, examined, and improved. They allow us to learn to collectively improve our understanding of the world and can be easily translated into quantitative simulation models.

The concepts introduced so far allow us to present a *first fundamental law* of systems thinking:

1. LAW OF STRUCTURE AND COMPONENT INTERACTION: *In order to understand and control the dynamics in the world it is necessary to identify the systemic structures that make up this world*.

   (a) On the one hand, the behavior of a variable depends on the system in which it is included.
   (b) On the other hand, the behavior of the entire system depends on both its *logical* and *technical* structures, that is, on the interconnections among its component variables and on the "machines" producing their variations.
       This law has *three corollaries*:

       • It is useless to try to modify the values of a variable if first we do not understand the systemic structure of which it is a part, since the balancing loops will restore its value and the reinforcing loops will increase it.
       • Even if we are not aware of it, in every systemic context the reinforcing loops are always linked to some balancing loop and vice versa.
       • In observing a dynamic world, the "*ceteris paribus*" assumption is never valid. Connected to the preceding fundamental law is a *second fundamental law* of systems thinking, which I shall name as:

2. LAW OF DYNAMIC INSTABILITY: *Expansion and equilibrium are processes that do not last forever; they are not propagated ad infinitum. Sooner or later stability is disturbed. Sooner or later the dynamics are stabilized* (see Sect. 1.6.3).

1.6   Complementary Material

1.6.1   Reinforcing Loop: Arms Escalation

Reinforcing loops represent many typical processes in which the loop variables reciprocally increase their values. This *section* presents basic reinforcing loops to represent escalation processes where X and Y continually reinforce each other. The subsequent *section* will consider a "ring" that shows a typical explosive process, in which a variable reinforces itself through increases generated by the variable itself.

To simulate a "ring" that represents an escalation process we go back to the model in Fig. 1.3, which describes the arms race, simulating this with equations (1.2) using Excel, whose use is particularly effective and simple to use for "small" loops formed by a limited number of variables (no more than eight or ten, in my experience). The simulation model that reveals the arms dynamics of two nations, X and Y, is illustrated in Fig. 1.10.

In the stylized loop that serves as a *control panel* we can immediately read the values assigned to the action and reaction rates as well as the initial values of the two variables. I have introduced the assumption that X, after observing the low initial level of Y's armaments (equal to $Y_0 = 50$), compared to its own level of $X_0 = 200$, overvalues its initial superiority. I also assume that X has a low action rate, $h(X/Y) = 10\%$, compared to an action rate for Y of $g(X/Y) = 15\%$, as well as a higher reaction time of $rX = 2$. This means that while Y reacts more energetically to variations in X's armaments, since $g = 15\% X$ reacts less intensely ($h = 10\%$) and more slowly ($r = 2$). The arms dynamics (Fig. 1.10) show that X's initial overestimation of its own strength allows Y to achieve a superior level of armaments starting already in the seventh year.

It is clear that the model does not explicitly represent the generating processes of the arms dynamics of X and Y. These dynamics are in fact produced when, due to processes fed by psychological and social motivations considered as a black box, increases in the armaments of one side, say X, increase the fear of Y's leaders and vice versa. This fear moves Y's leaders to increase their arms, but, following a typical escalation process, the increase in Y's arms causes an increase in X's arsenal; loop [**R**] acts over many periods. To take these variables into account we need a new model, which we shall examine in Sect. 1.6.4.

1.6.2   Reinforcing Loops: Rabbit Explosion

We know that rabbits are a big ecological problem in Australia. Rabbits were imported to Australia at the end of the eighteenth century (probably in 1788) to provide a supply of meat to the inhabitants. Around the middle of the nineteenth century 24 rabbits (along with other animals) were brought from England as prey for sport hunters. Rabbits imported, particularly strong and fast, crossbred with the

g(Y/X) | 15.00% |

r(X/Y)
| 1 |

arms X₀ | Variab. X | ... R ... | Variab. Y | arms Y₀
| 200.00 | | 50.00 |

| 2 |
r(Y/X)

| 10.00% | h(X/Y)

| time | Variab. X | | | Variab. Y | | |
	Arms X	rate X/Y	ΔY	Arms Y	rate Y/X	ΔX
0	200.0	15.0%	30.0	50.0	10.0%	2.5
1	202.5	15.0%	30.4	80.0	10.0%	4.0
2	206.5	15.0%	31.0	110.4	10.0%	5.5
3	212.0	15.0%	31.8	141.4	10.0%	7.1
4	219.1	15.0%	32.9	173.2	10.0%	8.7
5	227.7	15.0%	34.2	206.0	10.0%	10.3
6	238.0	15.0%	35.7	240.2	10.0%	24.0
7	262.1	15.0%	39.3	275.9	10.0%	27.6
8	289.7	15.0%	43.4	315.2	10.0%	31.5
9	321.2	15.0%	48.2	358.6	10.0%	35.9
10	357.0	15.0%	53.6	406.8	10.0%	40.7
11	397.7	15.0%	59.7	460.4	10.0%	46.0

[omitted]

Fig. 1.10 Arms escalation (see Fig. 1.3)

rabbits that were already there, causing them to quickly proliferate, setting off a typical *population explosion* phenomenon for rabbits. No attempt at control (hunting, destruction of warrens, infection of rabbits with the myxoma virus, etc.) was successful.

This population explosion for rabbits follows a typical reinforcing *Ring*, [**R**], which produces an exponential increase clearly shown in the model in Fig. 1.11, to construct which I have introduced the following assumptions (obviously the simulation model allows for these assumptions to be modified):

(a) A period of 100 years; an initial population of $X_0 = 24$ pairs; each of the newly imported specimens pairs with the one already present.
(b) A rate of population increase (fecundity rate–mortality rate) at the start of the process of $g_0 = 6$ %; it is assumed that, partly as a result of the attempts by authorities to limit the phenomenon, the rate "g" diminishes with the increase in population: $g_{50} = 3.5$ %; $g_{59} = 3$ %; $g_{75} = 2.5$ %; $g_{88} = 2$ %.
(c) It is assumed that for each generation only a percentage $h = 75$ % of the new individuals contributes to increasing the population; the remaining ones are prey for the predators; we assume that there are six pairings per year; this datum is inserted into the model by simply setting $r = 0.167$.

The model shows that after 100 years—the equivalent of 600 generations ($r = 0.167$, didactic assumption)—the rabbit population reaches about $X_{100} = 2.30$ billion pairs of rabbits. Since Australia has a surface area of $7.617.930 \text{ km}^2$, the model shows that after 100 iterations there will be an average population of 300 pairs of rabbits per square kilometer. If we assume that the authorities succeed in introducing new predators for rabbits, so that the number of newborns not yet at the reproductive age falls to $h = 50$ %, then the population would reach a density of only one pair per square kilometer.

I do not know if the data is correct and reasonable, but the model is presented for the sole purpose of illustrating the action of the loops [**R**] representing a typical population growth phenomenon; knowledge of the real birth and death rates of rabbits in Australia would allow us to refine the results of the model. Note that the population explosion model is valid for any type of "population" whose increase is exponential. Loop [**R**], which regulates this phenomenon, can also represent an annihilation process if the action rate "g" is negative and represents a rate of decrease that is self-reinforcing to the point that X is eliminated.

1.6.3   The General Law of Dynamic Instability

At the end of Sect. 1.5 I mentioned a general and intuitive law of systems thinking which I have called the:

g(Y/X) 6.00%

rX
1
R

rabbits₀ | Variab. X
24

0.16
rY

Variab. Y   New rabbits_t

75.00% h(X/Y)

| time | Variab. X | | | Variab. Y | | |
	rabbits0	g(Y/X)	ΔY	New rabbitst	h(X/Y)	ΔX
0	24.0	6.0%	9.0	9.0	75.0%	6.8
1	30.8	6.0%	11.5	11.5	75.0%	8.6
2	39.4	6.0%	14.8	14.8	75.0%	11.1
3	50.5	6.0%	18.9	18.9	75.0%	14.2
4	64.7	6.0%	24.3	24.3	75.0%	18.2
5	82.9	6.0%	31.1	31.1	75.0%	23.3
6	106.2	6.0%	39.8	39.8	75.0%	29.9
7	136.0	6.0%	51.0	51.0	75.0%	38.3
8	174.3	6.0%	65.4	65.4	75.0%	49.0
9	223.3	6.0%	83.7	83.7	75.0%	62.8
10	286.1	6.0%	107.3	107.3	75.0%	80.5
[omitted]						
89	860,072,331.0	2.0%	107,509,041.4	107,509,041.4	75.0%	80,631,781.0
90	940,704,112.1	2.0%	117,588,014.0	117,588,014.0	75.0%	88,191,010.5
91	1,028,895,122.6	2.0%	128,611,890.3	128,611,890.3	75.0%	96,458,917.7
92	1,125,354,040.3	2.0%	140,669,255.0	140,669,255.0	75.0%	105,501,941.3
93	1,230,855,981.6	2.0%	153,856,997.7	153,856,997.7	75.0%	115,392,748.3
94	1,346,248,729.9	2.0%	168,281,091.2	168,281,091.2	75.0%	126,210,818.4
95	1,472,459,548.3	2.0%	184,057,443.5	184,057,443.5	75.0%	138,043,082.7
96	1,610,502,631.0	2.0%	201,312,828.9	201,312,828.9	75.0%	150,984,621.7
97	1,761,487,252.6	2.0%	220,185,906.6	220,185,906.6	75.0%	165,139,429.9
98	1,926,626,682.6	2.0%	240,828,335.3	240,828,335.3	75.0%	180,621,251.5
99	2,107,247,934.0	2.0%	263,405,991.8	263,405,991.8	75.0%	197,554,493.8
100	2,304,802,427.9	2.0%	288,100,303.5	288,100,303.5	75.0%	216,075,227.6

- - -New Rabbits   ——rabbits

Fig. 1.11 Rabbit explosion

LAW OF DYNAMIC INSTABILITY: *Expansion and equilibrium are processes that do not last forever; they are not propagated ad infinitum. Sooner or later stability is disturbed. Sooner or later the dynamics are stabilized.*

Operationally speaking, this law affirms in a simple way that, though we are unaware or unable to observe this, every expansion loop is always associated with a
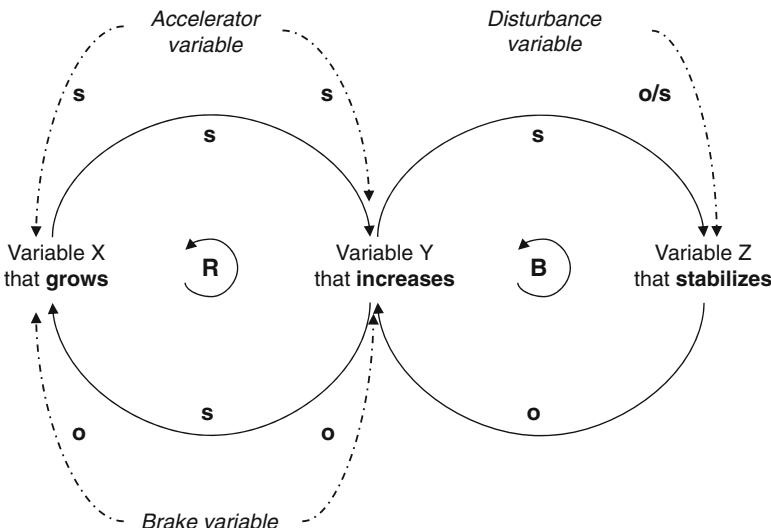
Fig. 1.12   The model representing the general law of dynamic instability

balancing loop that dampens the expansion dynamics; and vice versa, every balancing loop is associated with some type of expansion loop that counters the balancing effect. Moreover, external disturbances can come from the environment in the form of braking variables that counter the expansion or acceleration variables that eliminate the balancing effect, as shown in the model in Fig. 1.12.

> Paraphrasing Newton's first law of mechanics: "Every object remains in its state of rest or uniform motion in a straight line unless a force intervenes to modify this state", Systems Thinking could instead state: "Every repetitive system does not endlessly produce its own reinforcing or balancing processes because other processes intervene to reverse the dynamics". It seems impossible to respect the wise motto: "Quieta non movere, mota quietare!" (Mella 2012, p. 73).

A simple and interesting numerical simulation—in one of its many possible variants—is presented in Fig. 1.13. Several simulation tests are presented in Fig. 1.14. The diagram shows that the curve in bold, which depicts the dynamics of the stabilized variable Y_t, has a cyclical dynamics precisely due to the reinforcement from the curve for variable X_t and the balancing effects from the curve for variable Z_t.

By modifying the values in the control panel (directly shown in the CLD at the top of the table) we obtain the explosive dynamics of growth in Y if variable X prevails (see test 4 in Fig. 1.14) or the explosive decreasing dynamics if Z prevails. By introducing outside disturbances the dynamics can become quite irregular. I will not add any comments since the loops produced by the systems thinking technique are clearly representative and understandable. By adopting the simulation model behind Fig. 1.13 we can simulate other dynamics, introducing new parameters or even disturbance elements for one or even all three of the variables.

action rate Y/X 10.00%   action rate Z/X 20.00%

Reaction time Y 1   Reaction time Z 1

s   R   B   s

Variab. X X_0 100.00 Y_0 1,000.00 Variab. Y Z_0 400.00   Variab. Z

Reaction time X 1   Reaction time Y 1

s   o

12.00% action rate X/Y   15.00% action rate Y/Z

	Variab. X				Variab. Y			rate X/Y	ΔX	Variab. Z		
time	X	rate Y/X	ΔY		Y	rate Y/X	ΔZ	rate X/Y	ΔX	Z	rate Y/Z	ΔY
0	100.0	10.00%	10.0		1,010.0	20.00%	202.0	12.00%	121.2	602.0	15.00%	-
1	221.2	10.00%	22.1		941.8	20.00%	188.4	12.00%	113.0	790.4	15.00%	- 90.3
2	334.2	10.00%	33.4		856.7	20.00%	171.3	12.00%	102.8	961.7	15.00%	- 118.6
3	437.0	10.00%	43.7		756.1	20.00%	151.2	12.00%	90.7	1,112.9	15.00%	- 144.3
4	527.8	10.00%	52.8		642.0	20.00%	128.4	12.00%	77.0	1,241.3	15.00%	- 166.9
5	604.8	10.00%	60.5		516.3	20.00%	103.3	12.00%	62.0	1,344.6	15.00%	- 186.2
6	666.7	10.00%	66.7		381.2	20.00%	76.2	12.00%	45.7	1,420.8	15.00%	- 201.7
7	712.5	10.00%	71.2		239.4	20.00%	47.9	12.00%	28.7	1,468.7	15.00%	- 213.1
8	741.2	10.00%	74.1		93.2	20.00%	18.6	12.00%	11.2	1,487.3	15.00%	- 220.3
9	752.4	10.00%	75.2		- 54.7	20.00%	- 10.9	12.00%	-6.6	1,476.4	15.00%	- 223.1
10	745.8	10.00%	74.6		- 201.6	20.00%	- 40.3	12.00%	-24.2	1,436.1	15.00%	- 221.5
[omissis]												
91	- 1,821.8	10.00%	- 182.2		533.1	20.00%	106.6	12.00%	64.0	- 336.4	15.00%	66.5
92	- 1,757.8	10.00%	- 175.8		437.8	20.00%	87.6	12.00%	52.5	- 248.8	15.00%	50.5
93	- 1,705.3	10.00%	- 170.5		334.6	20.00%	66.9	12.00%	40.1	- 181.9	15.00%	37.3
94	- 1,665.2	10.00%	- 166.5		225.3	20.00%	45.1	12.00%	27.0	- 136.9	15.00%	27.3
95	- 1,638.1	10.00%	- 163.8		112.1	20.00%	22.4	12.00%	13.4	- 114.5	15.00%	20.5
96	- 1,624.7	10.00%	- 162.5		3.2	20.00%	0.6	12.00%	-0.4	- 115.1	15.00%	17.2
97	- 1,625.1	10.00%	- 162.5		- 118.5	20.00%	- 23.7	12.00%	-14.2	- 138.8	15.00%	17.3
98	- 1,639.3	10.00%	- 163.9		- 231.6	20.00%	- 46.3	12.00%	-27.8	- 185.1	15.00%	20.8
99	- 1,667.1	10.00%	- 166.7		- 340.5	20.00%	- 68.1	12.00%	-40.9	- 253.2	15.00%	27.8
100	- 1,707.9	10.00%	- 170.8		- 443.3	20.00%	- 88.7	12.00%	-53.2	- 341.9	15.00%	38.0

Fig. 1.13 The general law of dynamic instability: A simulation

The first three tests reveal the relative stability in the dynamics of variable Y (the thicker curve), which is stabilized even in the presence of disturbances; in fact, the latter are "neutralized" by the joint action of the variables X and Z. Test 4 shows that by deactivating loop [**B**] we get only the reinforcing effect of loop [**R**].

1.6.4 The Law of Dynamic Instability: Richardson's Model

An initial example of how we can apply the law of dynamic instability to give a more correct representation of dynamic systems comes from reformulating the arms escalation model in Fig. 1.10, which only contained a single loop [**R**] that produced an apparently infinite growth in X's and Y's arms.

To avoid this infinite growth, which is entirely unrealistic, we must introduce, in addition to the *action* and *reaction rates* "g" and "h," respectively—which we can call "coefficients of defense," since they indicate to what extent fear moves X and Y to increase their arms—two new coefficients as well, called the "coefficients of saturation," since they express the degree of tolerability in the population for the economic sacrifices it must make to strengthen their country's armaments. We then get the model in Fig. 1.15, which adds two loops [**B**] to loop [**R**]; these loops, as stated in the law of dynamic instability, impede an uncontrolled growth in arms.

Test 1	
X0 =	100
Y0 =	1.000
Z0 =	400
g(Y/X)=	10%
g(Z/Y)=	20%
g(Y/Z)=	15%
g(X/Y)=	12%
r(X) =	1
r(Y) =	1
r(Z) =	1
NO disturbances	

Test 2	
X_0=	100
Y0=	1.000
Z0=	400
g(Y/X)=	10%
g(Z/Y)=	20%
g(Y/Z)=	15%
g(X/Y)=	12%
r(X)=	1
r(Y)=	1
r(Z)=	1
YES disturbances	

Test 3	
X_0 =	1.000
Y_0 =	1.000
Z_0 =	1.000
g(Y/X) =	30%
g(Z/Y) =	30%
g(Y/Z) =	15%
g(X/Y) =	20%
r(X) =	1
r(Y) =	1
r(Z) =	1
YES disturbances	

Test 4	
X_0 =	1.000
Y_0 =	1.000
Z_0 =	1.000
g(Y/X)=	10%
g(Z/Y)=	0%
g(Y/Z)=	0%
g(X/Y)=	3%
r(X)=	1
r(Y)=	1
r(Z)=	1
NO disturbances	
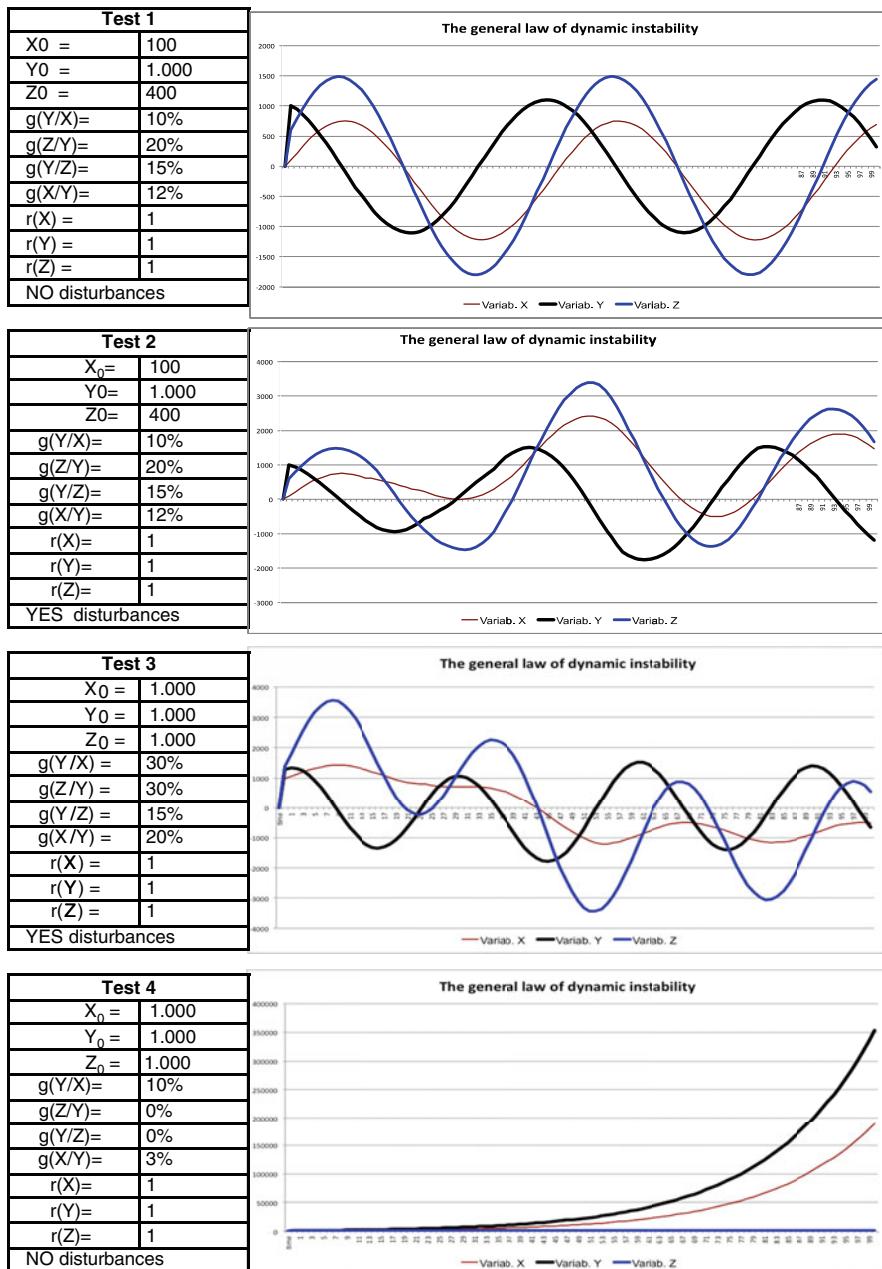
Fig. 1.14 The general law of dynamic instability: Simulation test (the *bold line* represents the stabilized variable *Y*)

	Variable X			Variable Y						
years	Arms X	def. C. Y	ΔY	Arms Y	def. C. X	ΔX	sat. R. X	ΔX	sat. R. Y	ΔY
0	500.0	15.0%	75.0	175.0	5.0%	8.8	-15.0%	- 75.0	-5.0%	- 8.8
1	433.8	15.0%	65.1	231.3	5.0%	11.6	-15.0%	- 65.1	-5.0%	- 11.6
2	380.3	15.0%	57.0	276.8	5.0%	13.8	-15.0%	- 57.0	-5.0%	- 13.8
3	337.1	15.0%	50.6	313.5	5.0%	15.7	-15.0%	- 50.6	-5.0%	- 15.7
4	302.2	15.0%	45.3	343.2	10.0%	34.3	-10.0%	- 30.2	-5.0%	- 17.2
5	306.3	15.0%	45.9	371.9	10.0%	37.2	-10.0%	- 30.6	-5.0%	- 18.6
[omitted]										
20	973.6	20.0%	194.7	1,355.6	10.0%	135.6	-5.0%	- 48.7	-5.0%	- 67.8
21	1,060.5	20.0%	212.1	1,499.9	10.0%	150.0	-5.0%	- 53.0	-5.0%	- 75.0
22	1,157.4	10.0%	115.7	1,540.6	5.0%	77.0	-5.0%	- 57.9	-5.0%	- 77.0
23	1,176.6	10.0%	117.7	1,581.2	5.0%	79.1	-5.0%	- 58.8	-5.0%	- 79.1
24	1,196.8	10.0%	119.7	1,621.9	5.0%	81.1	-5.0%	- 59.8	-5.0%	- 81.1
25	1,218.1	10.0%	121.8	1,662.6	5.0%	83.1	-5.0%	- 60.9	-5.0%	- 83.1
[omitted]										
56	1,223.8	10.0%	122.4	1,492.5	10.0%	149.2	-15.0%	- 183.6	-10.0%	- 149.2
57	1,189.5	10.0%	118.9	1,462.2	10.0%	146.2	-15.0%	- 178.4	-10.0%	- 146.2
58	1,157.3	10.0%	115.7	1,431.7	10.0%	143.2	-15.0%	- 173.6	-10.0%	- 143.2
59	1,126.8	10.0%	112.7	1,401.2	10.0%	140.1	-15.0%	- 169.0	-10.0%	- 140.1
60	1,097.9	10.0%	109.8	1,370.9	10.0%	137.1	-15.0%	- 164.7	-10.0%	- 137.1

Fig. 1.15 Richardson's model

Figure 1.15 is based on the model by Lewis Richardson (1949), considered to be the founder of the scientific analysis of conflicts.

Not taking into account, for simplicity's sake, other variables introduced by Richardson, I have reformulated the model by assuming that *Y*'s arms increase due

to fear of the size of X's arsenal, according to a (positive) *coefficient of defense*, and undergo a slowing down proportionate to the size of the armaments themselves, according to a (negative) *coefficient of saturation*.

Since X starts with an arsenal five times that of Y's ($X_0 = 500$ against $Y_0 = 100$), I have set a low *coefficient of defense* ($h = 5\%$ against $g = 15\%$ for Y) as well as a high *rate of saturation* (15% against 5% for Y), since we can assume that the population of X is intolerant toward additional sacrifices to increase its already high arms level. On the other hand, Y's citizens, conscious of their initial inferiority, are very frightened by X's arsenal and will tolerate even substantial sacrifices to shore up their security. We can observe these dynamics in many contexts, even in the recent developments in the conflicts in the Middle and Far East.

The model in Fig. 1.15 is useful for understanding the opposite phenomenon of a gradual reduction in recent years in the nuclear arms of the two greatest world powers, which has occurred due to both the increase in the saturation rate and the decrease in the defense coefficient. This tendency has continued even recently, and according to what was published in America24 on March 24, 2010:

> Washington and Moscow are close to signing a new nuclear treaty. U.S. president Barack Obama briefed members of Congress on the state of negotiations with Russia on the reduction of nuclear arms. All aspects of the agreements are said to have been defined and the formal announcement could come in the next few days. Negotiations between American and Russian representatives have been going on for months with the aim of signing a new treaty to replace the 1991 treaty, which expired in December 2009 (www.america24.com).

1.6.5   The Law of Dynamic Instability: Prey–Predator Populations

Section 1.6.2 presented a loop [**R**] to simulate population growth. However, the model in Fig. 1.11 is unrealistic and openly contradicts the law of dynamic instability. No population grows infinitely, since growth is always balanced by at least three factors operating individually or jointly: a maximum limit of available resources and the presence of predators and human controls. I propose the very simple model in Fig. 1.16, which represents the dynamics of two populations, sardines and sharks, which, being preys and predators, limit in turn their growth.

The more sardines increase in number, the greater the increase in the shark population due to the abundance of food; this increase in shark numbers reduces the number of sardines, which causes a reduction in the number of predators. These dynamics are described by a balancing loop that produces oscillating dynamics for the two populations normally limited in size. I would note that the study of population dynamics has a solid tradition going back to Vito Volterra and Alfred Lotka and to their famous equations that describe the coevolution of prey–predator populations. Note that the model in Fig. 1.16 does not derive from these equations but is presented with the sole aim of illustrating in a simplified way how the law of dynamic instability works (the original equations are examined in Sect. 7.2).

c(Y/X) =   0.10

0
a - accelerator

40
b - accelerator

s   r sharks =   1

X   **B**   **Y**
SARDINES   SHARKS

2,000
X(t = 0)

r sardines =   1   **o**

300
Y(t = 0)

Δx(t=0) =

2%

| 5.00 | = cX/Y)
|---|

	SARDINES		**SHARKS**	
Periodi = t	SARDINES	ΔX	SHARKS	ΔY
0	2,000.0	40.0	300.0	44.0
1	2,040.0	- 179.2	344.0	22.1
2	1,860.8	- 73.2	366.1	32.7
3	1,787.6	- 127.7	398.8	27.2
4	1,660.0	- 103.0	426.0	29.7
5	1,557.0	- 117.4	455.7	28.3
6	1,439.6	- 112.5	484.0	28.7
7	1,327.1	- 117.2	512.7	28.3
8	1,209.9	- 117.2	541.0	28.3
9	1,092.7	- 119.5	569.3	28.0
10	973.1	- 120.8	597.3	27.9

Fig. 1.16 Prey–predator dynamics (the model is a very simplified application of the Volterra–Lotka equations)

1.6.6   Systems Archetypes: Three Relevant Structures of Human Behavior

A particularly useful class of systems that systems thinking deals with are *systems archetypes*, or standard structures, that is, *simple general* and *stable* models of relations that frequently recur in various situations in any type of organization, public as well as private companies, and in different environments.

> One of the most important, and potentially more empowering, insights to come from the young field of systems thinking is that certain patterns of structure recur again and again. This "systems archetypes" or "generic structures" embody the key to learning to see structures in our personal and organizational lives. [...] Because they are subtle, when the archetypes arise in a family, an ecosystem, a news story, or a corporation, you often don't see them as so much as feel them. Sometimes they produce a sense of deja vu, a hunch that you've seen this pattern of forces before. "There it is again" you say to yourself (Senge 2006, p. 93).

The aim of archetypes is to rapidly increase the capacity of the observer, the manager, the governor, or the decision-maker to see the systemic problems, recognize the structures that determine these problems (Sect. 5.9), and identify leverage effect of these structures for the purpose of formulating definitive solutions. Senge (1990) describes ten archetypes that have been adopted by all of his followers and which by now have become part of the shared heritage of knowledge of this field. I hold that, along with Senge's ten standard archetypes, we can identify other archetypes, equally important and widespread (Mella 2012), several of which we shall examine in subsequent chapters (Chap. 5, in particular).

In my opinion there are three basic *archetypes*, which usually occur together and act in a unitary manner, which account for a vast variety of behaviors by human agents and which also justify the validity of many of Senge's standard archetypes. These basic structures describe man's natural instinct, reinforced by the prevailing educational and cultural models, not to give enough importance to the harm his behavior can produce; as a result, in the face of such harmful consequences the human agents refrain from modifying their cultural attitudes and applying systems thinking to assess the disastrous effects of their behavior.

The first archetype, which I propose to call the "short-term preference archetype" (Mella 2012, p. 234), is innate in all men who adopt a "shortsighted" rationality principle. This archetype manifests itself when a particular behavior is repeated over time and produces, *at the same time*, immediate, *short-term advantages* (benefits, or pleasures) and *disadvantages* (sacrifices, or harm) that occur over the *long run* with a significant delay. Here it is not a question of choosing between two or more advantages or two or more disadvantages, distributed differently over time, but of understanding how an immediate, or short-term, certain *advantage* is evaluated when accompanied by a long-term, probable, even notable *disadvantage* under the assumption that these are produced by the *same repetitive behavior*.

Since the short-term *advantages* are immediately perceived, it is "natural" to attribute high value to them, and this leads the agents to prefer and repeat the behavior that produces them. At the same time, the perception of the short-term
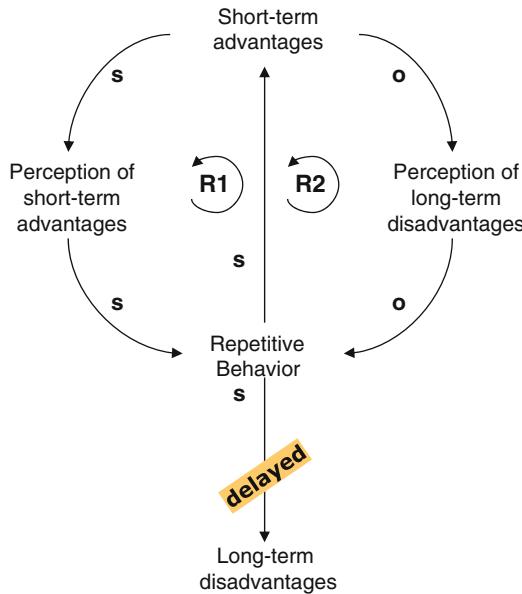
Fig. 1.17   Archetype of short-term preference (*source*: Mella 2012, p. 235)

advantages is accompanied by a diminished perception (probability, weight, judgment) of the long-term *disadvantages*, which makes it likely the agent will repeat the behavior that produces the short-term advantages, even if long-term disadvantages are inexorably associated with these. Figure 1.17 shows the simple structure of the "short-term preference archetype": the repetitive behavior produces short-term advantages (vertical arrow "**s**") which, in turn, encourage the continuation of the behavior, according to loop [**R1**].

The current advantages reduce the perception of the long-term disadvantages—which are thus produced with a considerable *delay*—and this encourages even more of the same behavior in order to gain present advantages, as shown in loop [**R2**], so that the long-term disadvantages do not condition behavior but represent only its effect (no loop).

The two loops act together to increase *short-term preferences*, which guide behavior and inexorably create long-term problems when the *disadvantages* caused by the repeated behavior appear. It is useful to point out that the terms *advantages* and *disadvantages* must be taken broadly. The chance to avoid a disadvantage (harm, an unpleasant event) could be perceived as an advantage, and a missed advantage could be perceived as a disadvantage. In any case, the structure in Fig. 1.17 would not change much. The *short-term* advantage prevails over the *long-term* disadvantage in conditioning behavior. There is no way out: this archetype is always in action, and when it is not it is always lurking.

Let us reflect on the behavior of parents who prefer to overfeed their children, allow them to watch TV, or play with the PC for hours and hours in order to avoid any disturbance or complaint today, not thinking about the long-term physical and
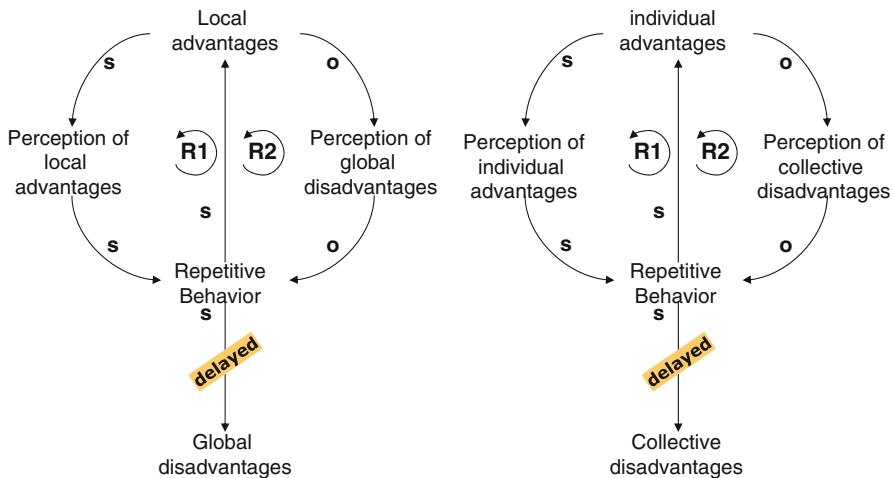
Fig. 1.18 Archetype of local (*left*) and individual (*right*) preference (*source*: Mella 2012, p. 236)

psychological consequences. Consider pollution "today" and the lack of interest in the damage "tomorrow." The behavior of the smoker who prefers the short-term pleasure of thousands of cigarettes to the danger from future harm, perceived as remote, is exactly the same as that of the person who prefers a convenient disposal of waste, accumulating this in landfills or burning it in incinerators, thereby demonstrating a lack of interest in the damage "tomorrow." Let me take another example: spending cuts for present academic research or health care to favor a balanced budget "today" will cause relevant damage to the state of "medium- to long-term" competitiveness and health status of the population. There is a popular saying: "The ant is always wiser than the cicada," but the cicada prefers "today" the pleasure that "tomorrow" will lead to its death. Not to mention the behavior of the careless driver . . ..

The leverage effect that derives from this ARCHETYPE is fundamental: when we perceive its action it is useless to try and oppose loop [**R1**] (anti-smoking laws, emission standards); instead, it is necessary to weaken loop [**R2**] through a convincing information and training campaign, day after day, to make the decision-maker aware of the disadvantages the behavior can cause (awareness campaign, company training). The lack of respect for the Kyoto protocol is mainly to be attributed to those countries that do not want to limit their emissions in order not to undermine growth "today," not worrying about the damage global warming will cause in upcoming decades.

The *archetype* of short-term preference has two parallel variants that are equally widespread and ruinous:

1. The agent prefers the objectives that bring *local* advantages over those that would provide *global* ones; this is the "local preference archetype" (Fig. 1.18 left).
2. The agent prefers the objectives that produce *individual* advantages over those for larger *groups*; this is the "individual preference archetype (Fig. 1.18 right).
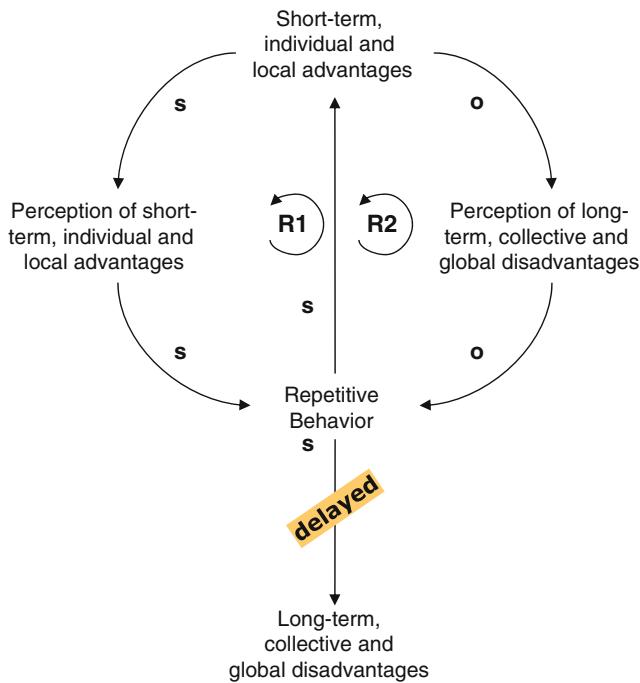
Fig. 1.19 Archetype of short-term, local, and individual preference

Because the local and "partisan" advantages are perceived in the short term while the global and collective ones are usually noted in the long term, these structures can be considered as particular cases of the preceding one. Nevertheless, we can distinguish them by referring to them as the *archetypes* of "individual preference" (Mella, 2012: 236).

We are all inclined to give priority to objectives that bring advantages to our family, club, or association rather than to those that bring advantages to the collectivity. Anyone working in a department tries to make it function in the best way possible, even if this conflicts with the functioning of other departments. Judgments on the local disadvantages from public works (rubbish dumps, incinerators, dams, highways, etc.) prevail over those regarding the global advantages they will provide the larger territory.

Even if we are aware that public landfills or incinerators are necessary for the common good, we steadfastly refuse to allow these to be placed in our towns, since local interests prevail over the global advantages they will provide the larger territory. The NIMBY effect is always present. "Not in my back yard" is the most common individual response to the community's request to put up with the construction of public works for the advantage of the collectivity.

It is easy to see that the three archetypes often overlap; thus, it is useful to present them using a single model, as shown in Fig. 1.19.
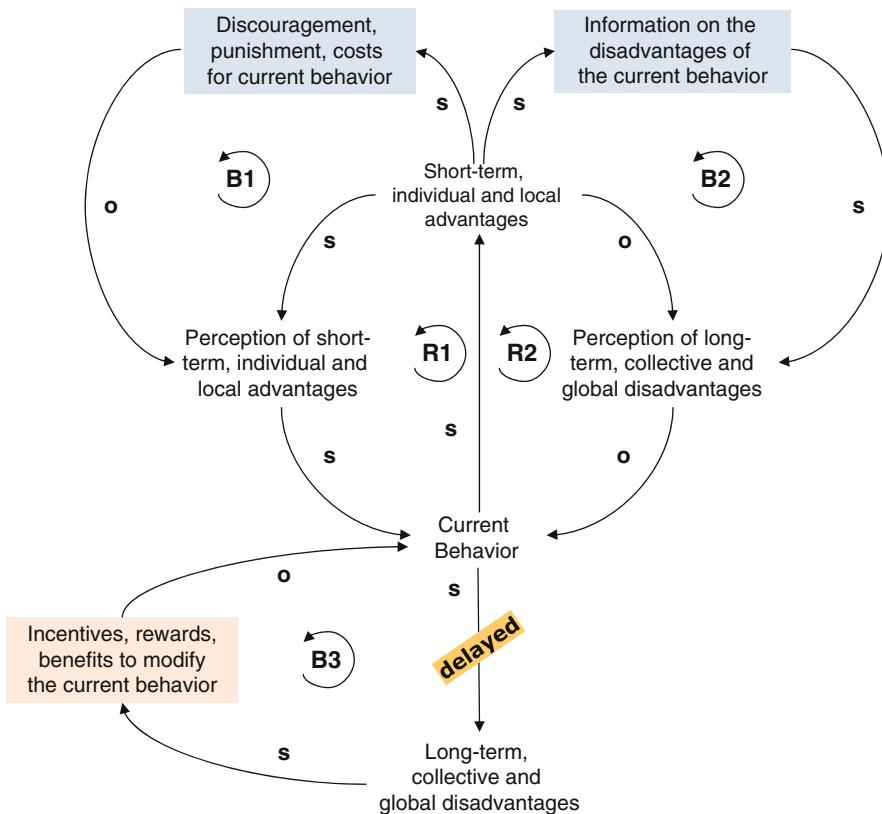
Fig. 1.20 Actions to weaken the behavioral archetypes

The three archetypes that dominate the prevailing form of individual–agent behavior are very powerful, since the individual that carries out the archetypes sees them as entirely rational. In fact, every agent naturally behaves based on his or her own individual, local, and short-term rationality, not "willingly" accepting being guided by a global rationality. The short-term, local, and individual advantage prevails over the long-term, global, and collective one in conditioning behavior. Therefore, if we wish to avoid the harmful consequences of the actions of the three archetypes it is necessary to act to counter the reinforcing effect of loops [**R1**] and [**R2**]. There are three important weakening actions for such loops; to better understand these we can refer to Fig. 1.20, where the three archetypes have been grouped into a single model of equal significance.

The first solution is to try to weaken or eliminate loop [**R1**]. Since the local advantages would immediately reveal the benefit of the individual action, an appropriate system of *disincentives*, even in the form of sanctions or the imposition of costs and taxes, proportional in some way to the amount of short-term local advantages, would lessen the perception of the convenience from repeating the

current behavior, thus forming the balancing loop [**B1**], which could compensate the effects of loop [**R1**].

A second way to counter the archetype is to weaken loop [**R2**], which produces an underestimate by the agent of the long-run disadvantages. An incisive, widespread, and continual *stream of information* about the negative long-run global effects due to short-term, local, and individualistic behavior would lead to a conscious awareness by the agent of the global harm from his or her behavior. The persuasive capacity of this *information* must be proportional to the amount of short-term local advantages. The more relevant these are, the more the information must be insistent, repeated, uniform, credible, convincing, and diffused over vast territories.

The third form of intervention to reduce the negative effects from the action of the archetype in Fig. 1.19 is to act directly on the agent's current behavior through incentives, [**B3**], even in the form of social and economic *benefits*, rewards, tax breaks, etc., to induce the agent to modify his or her current behavior by reducing its intensity and frequency, adopting processes and technologies that reduce the long-term global disadvantages.

The modus operandi of the three archetypes will be taken up below in Sects. 4.7 and 6.7.

1.6.7 The "Six" Disciplines of Learning Organizations

Peter Senge introduced systems thinking not as a new theory of systems but as the fifth of the disciplines needed to construct a *learning organization*.

Referring the interested reader to Senge's work and to the vast literature on the topic (Mella 2012), here I will merely summarize Senge's theory for those who do not need to study it in greater detail.

The first discipline is "personal mastery." In order to improve the organization's performance it is necessary to develop the discipline of personal mastery, which, bringing out in each individual of the organization the desire to learn and create, teaches him or her to face life in a creative and nonreactive way, which favors individual growth through personal learning and improvement, thereby contributing to clarifying the individual's view about his life and work.

> Personal mastery is the discipline of continually clarifying and deepening our personal vision, of focusing our energies, of developing patience, and of seeing reality objectively. As such, it is an essential cornerstone of the learning organization—the learning organization's spiritual foundation (Senge 2006, p. 7).

> The discipline of personal mastery [. . .] starts with clarifying the things that really matter to us, of living our lives in the service of our highest aspirations (*ibidem*, p. 8).

The second discipline, referred to as "mental models," is particularly relevant because it connects with the ability to clarify and improve the frames and repertoires that affect attention and decision-making and thus to ameliorate the action individuals undertake in their organization. In fact, a mental model is a "pattern" or

a "theory" that guides a person in the decisions and choices he or she makes when he or she has to act. If the results of the model are positive, the model is strengthened; otherwise, it is "set aside" and another model of behavior is sought (I shall deal more amply with the mental models in Sect. 9.14).

> 'Mental models' are deeply ingrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action (Senge 2006 p. 8).

Mental models influence in a crucial way the quality of thinking of individuals, but they are also pervasive in their effect on organizations, in the form of operational procedures (I do this because it is written so in the manual), widely accepted organizational practices (I act like this because everyone else does), standardized decisional rules (it might be wrong, but that is what I have been told to do), and so on. The mental models discipline is fundamental for organizational learning since it not only increases the group's or the individual's capacity to form a stock of shared knowledge but also facilitates the process for recognizing and modifying the group mental models in order to collectively decide in an effective way, as if the decision came from a single individual.

> The discipline of working with mental models starts with turning the mirror inward; learning to unearth our internal pictures of the world, to bring them to the surface and hold them rigorously to scrutiny (Senge 2006, p. 8).

"Building shared vision" is the third discipline. In order to build a shared vision the heads and leaders must continually analyze and share their personal visions with the other members of the organization and have these accepted by the latter, who then, in turn, spread them through personal commitment. As Senge observes, in most modern-day organizations relatively few people can be said to be enrolled and even a lesser amount to be committed. Most are in a state of conformism, whether formal or genuine. The enrolled people assume the role of gaining the attention of other members in order to improve their attention and knowledge. The discipline of creating a shared vision has always been fundamental for the long-term survival of organizations, but it is even more essential today for building learning organizations.

> If any one idea about leadership has inspired organizations for thousands of years, it's the capacity to hold a shared picture of the future we seek to create. [...] The practice of shared vision involves the skills of unearthing shared 'pictures of the future' that foster genuine commitment and enrollment rather than compliance (Senge 2006, p. 9).

The fourth discipline is referred to as "team learning," that is to say, the process that seeks to create and develop the team's capacity to focus the attention of all participants toward stimuli, categories, and objectives in order to work together in a coordinated manner to obtain results its members truly desire, perhaps in order to achieve a shared vision through dialogue, in order to listen to the different points of view, and discussion, in order to search for the best point of view to support the decisions that must be made.

> The discipline of team learning starts with "dialogue," the capacity of members of a team to suspend assumptions and enter into a genuine "thinking together." [. . .] "The discipline of dialogue also involves learning how to recognize the patterns of interaction in teams that undermine learning" (Senge 2006, p. 10).

Alignment is a necessary condition in order that the power given to an individual can increase the power of the entire team.

> [. . .] there are striking examples where the intelligence of the team exceeds the intelligence of the individuals in the team, and where teams develop extraordinary capacities for coordinated action (Senge 2006, p. 9).

The fifth discipline is "systems thinking." In order to develop a learning organization the four disciplines must, according to Senge, be jointly and coherently applied. This requires a fifth discipline, which serves as a unifier and coordinator for the other four, and it is this that justifies the role of systems thinking.

The role of the sixth discipline, "control thinking," is also clear. If learning organizations increase their survival capacity to the extent they are able to successfully react to environmental disturbances and produce a change in their position to achieve states more favorable to their survival—if the capacity to react and change depends on the extent to which they learn quickly as unitary systems, thus making the network of their processes even more efficient—then it is clear that they learn to the extent to which they can set their own objectives, translate these into coherent and shared individual objectives, verify their achievement, and take the necessary collective and individual actions to identify and eliminate deviations from the achievement of the objectives. As the following chapter will show, this is precisely the logic behind control.

The role of the sixth discipline is to contribute to the construction of learning organizations by promoting an education and discipline of control at all levels in order to transform the organizations themselves into networks of control systems, as I shall try to demonstrate in subsequent chapters.

1.7 Summary

Systems thinking—which Peter Senge considers the fifth discipline for forming learning organizations—represents an effective language and a powerful logical tool for understanding and describing dynamic systems as well as control systems, as we shall see in the next chapter (Sect. 1.1).

There are five basic rules of systems thinking (Sect. 1.2).

1. If we want to understand the world we must be able to "see the trees and the forest"; we must develop the capacity to "zoom" from the whole to the parts, from systems to components, and vice versa.
2. We must not limit our observation to that which appears constant but "search for what varies"; the variables are what interest the systems thinker.

3. If we truly wish to understand reality and change, we must make an effort "to understand the cause of the variations in the variables we observe."
4. It is not enough to search for the causes of the variations we observe only by searching for the causal chains among the variables; we must also "link together the variables in order to specify the loops among all the variations."
5. When we observe the world we must "always specify the boundaries of the system we wish to investigate."

The simplest models that describe dynamic systems are the reinforcing loops [**R**] and balancing loops [**B**]. Dynamic systems of whatever size are constructed by joining together an appropriate number of loops [**R**] and [**B**]. The basic balancing or reinforcing loops that connect two variables can also be called "*Rings*" since they usually assume a circular representation (Sect. 1.3).

Dynamic systems are repetitive systems equipped with memory and closed with respect to the environment. They can be simulated using system dynamics techniques (Sect. 1.4).

Two important laws of Systems Thinking have been presented (Sect. 1.5) along with three fundamental "archetypes" describing the human behavioral trait of preferring short-term benefits, even if these produce long-term damage (Sect. 1.6.6).

Chapter 2
The Ring. The General Structure of Control Systems

All processes that are stable we shall predict. All processes that are unstable we shall control... This was John von Neumann's dream (John von Neumann's dream, from Dyson 1988, p. 182).

In case a system approaches a stationary state, changes occurring may be expressed not only in terms of actual conditions, but also in terms of the distance from the equilibrium state; the system seems to "aim" at an equilibrium to be reached only in the future. Or else, the happenings may be expressed as depending on a future final state (Ludwig von Bertalanffy 1968, p. 75).

Stability is commonly thought of as desirable, for its presence enables the system to combine flexibility and activity in performance with something of permanence. Behaviour that is goal-seeking is an example of behaviour that is stable around a state of equilibrium. Nevertheless, stability is not always good, for a system may persist in returning to some state that, for other reasons, is considered undesirable (Ross Ashby 1957, p. 61).

Chapter 1 presented the principles of systems thinking along with the rules and formal language, and this discipline uses to represent dynamic systems. This chapter deals with the feedback control problem as a general form of efficient control. By adopting the systems thinking language we shall construct the general *logical model* of a *closed-loop control system*, the "Ring," which allows the process control to take place (Sects. 2.1–2.8). The feedback control system model represents the general foundation for every ordered and stable dynamic in the world, and thus the fundamental instrument by which the world itself maintains its existence. This approach allows us to develop a true "art" of control; a *discipline* based on systems thinking logic that enables us to observe, understand, and modify the world and our future. I shall define the *logical structure* of the *Ring* along with the *recursive equations* that determine its dynamics. The general *logical structure* of control systems is integrated with its *technical structure*, which is made up of four apparatuses or machines—effector, sensor, regulator, and information network—which, in fact, make up the *chain of control* (Sect. 2.9). The model also includes

three types of *delays* and *disturbances*—action, detection, and regulation—which characterize the processes carried out by the apparatuses. Other new and important concepts are introduced: (1) the management and governance processes; (2) the interferences between control systems; and (3) different areas of application of the general model.

2.1 The Control Process

As I have pointed out in Chap. 1, systems thinking views the world as a system of systems of interconnected variables, thereby constructing a "world of interacting variables" linked by causal chains, by reinforcing and balancing loops. Since the world is made up of variables, we might expect incredible disorder; variables that grow and diminish and that, through their dynamics, impact other variables, which in turn are connected to others, as part of a vortex of dynamics that combine in a disorderly fashion, as if we were observing the turbulent dynamics of the sun's surface, the atmosphere, or the oceans. Yet the world normally appears ordered, except for some exceptions in the context of complex systems (CSS 2013); the dynamics of variables seems regulated over time and in space, and any excessive variation is brought back to normality through special systems known as control systems.

In fact, control systems are systems that control the world's fundamental variables. Cells, as well as all living beings, can remain alive for a long time thanks to the order that stabilizes the multiple interactions among their constituent variables. If the disturbances are not so intense as to be destructive, then autopoiesis can continue over time, allowing the living being to exist. It is clear that when the temperature falls below zero men normally do not die of exposure, and that when the climate turns hot they do not die scorched and dehydrated. They are able to maintain their temperature within the standard vital parameters, except for very extreme variations in the outside temperature. If it is cold one dresses more warmly, lights a fire, lives in the barn with other animals, finds refuge in a cave, builds earthen houses that insulate better, even adopts modern temperature regulators and solar paneling, modern insulating fabric, air conditioners, and constructs thermoregulated skyscrapers.

What is true for intelligent man is also true for other species: winter animals go into hiding, hibernate or gather together in small herds. Penguins resist polar temperatures by forming large compact flocks and allowing the outside members to rotate from the outside to the inside of the flock to allow every member to be protected from the cold. In summer, animals adopt strategies against the heat, for example, by looking for watercourses or going under the sand. Evolution has provided many ways to regulate temperature, equipping animals with fur and layers of fat, even developing cold-blooded animals. The vegetable kingdom also has a wide variety of strategies against excessive heat or cold.

As subsequent chapters provide a wide range of examples, these few examples above will suffice to show the importance of control and control systems. Though there are several definitions of control and control systems, it is useful to start with the simple and significant one by Michael Arbib:
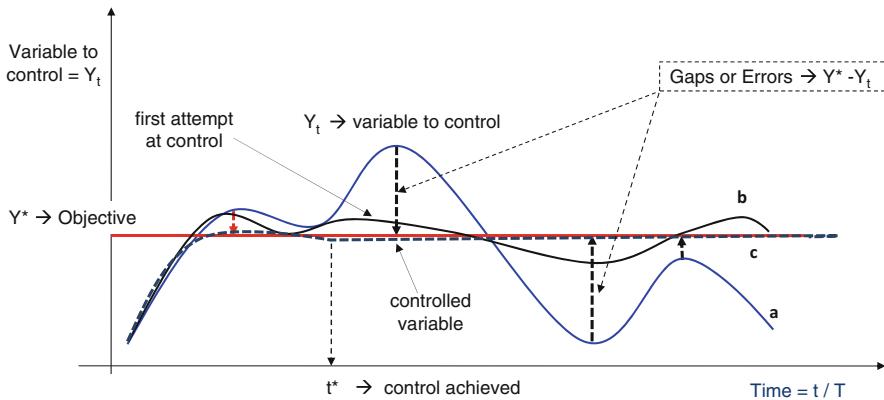
Fig. 2.1 Control problem

> In general terms, therefore, a control problem is to choose the input to some system in such a way as to cause its output to behave in some desired way, whether to stay near a set reference value (the regulator problem), or to follow close upon some desired trajectory (the tracking problem) (Arbib 1987, online).

Starting from the above quote, I shall define as "controllable" with respect to time a variable, Y_t, if it can take on in a succession of periods, or instants, $t = 1$, $2, \ldots$ (based on a temporal scale to be specified) a desired or established value, Y^*. The instant t^* in which the objective is achieved defines the *control instant*.

The value Y^* that the variable Y_t must attain can represent a *constraint* (for example, do not go below 10,000 ft; do not exceed 50 km per hour; emissions not to exceed 150 parts per million; costs not to exceed \$ 120 per unit of product), a *limit* (break the high jump record; the temperature of dew at which fog forms; a body temperature of 36 °C), an *objective* or *goal* (raise the temperature to 24 °C; reach 12,000 m of altitude; turn left, etc.). The distinction between *constraints*, *limits*, and *objectives* is not always clear; for this reason, unless clearly indicated I will use Y^* (set-point) to indicate an *objective* to be reached (as well as a *constraint* or a *limit*).

The graph in Fig. 2.1 clearly represents the control problem for a variable Y_t (curve a) whose dynamics must converge toward the objective Y^*, indicated by a half line and assumed to remain constant over time.

When the values for Y_t move away from those for Y^*, *gaps* develop which must be eliminated. The *gap* (Δ) is calculated as the difference $\Delta(Y)_t = Y^* - Y_t$, so that it is immediately clear that if $\Delta(Y)_t < 0$, then the dynamics of Y_t must be reduced, and vice versa if $\Delta(Y)_t > 0$. The *gap* can be given various names (deviation, difference, etc.), but henceforth I shall use the general term "error," indicating this by $E(Y)_t$, which is the equivalent of $\Delta(Y)_t$. An attempt at control Y_t noticeably reduces the error (curve b), though not entirely. The control is successful when the dynamics of Y_t at instant t^* coincide with the value Y^* (dashed line c) and $\Delta(Y)_t = 0$.

We must specify how it is possible to act on the variable Y_t so that, at successive instants t_1, t_2, t_3, etc., we can modify its values and make these move toward Y^*. The answer is clear: to control Y_t we always require the action of a control system.

2.2 The Logical Structure of Control Systems in the Language of Systems Thinking

After these introductory comments and with the aid of Fig. 2.1, we can easily introduce the concept of a control system by adopting the logic and language of systems thinking.

We know from systems thinking that the dynamics of Y_t can be considered the effect of one or more variables, X_t, that cause those dynamics. I shall define a *control variable* (*action* or *active* variable) as any variable X_t on whose values those of Y_t depend based on the causal relation, so that the variations in Y_t can be considered produced by variations in X_t.

For this reason, the variable X_t can also be referred to with the more evocative term *control lever*. The values assigned to X_t in successive instants are not "arbitrary" but are *calculated* and *readjusted* to *gradually* eliminate the error between X^* and the values of Y_t. I shall define as *manager* of the system the person who can calculate and modify the values of X_t; that is, in figurative language, "activate the control lever." This simple logic is translated into the *standard model* of Fig. 2.2, which shows the *logical structure* of the *single-lever* and *single-objective* control system, *without delays*, able to control only a single variable, as shown in Fig. 2.1 (*multi-lever* and *multi-objective* control systems will be examined in Chap. 4).

Without discussing for now the internal operating mechanisms (these will be taken up in Sect. 2.9), I shall define as a control system any *logical* or *technical* structure (algorithm or machine, rule or apparatus, etc.) which, for a set of instants, perceives $E(Y)_t = Y^* - Y_t$, calculates and assigns the values of the lever X_t, and produces the appropriate Y_t to gradually eliminate, when possible, the error $E(Y)$, thereby achieving the equality $Y_{t*} = Y^*$ at instant t^*.

Since the system in fact "perceives and evaluates" the errors with regard to the objective, it operates according to the logic referred to as *feedback control*.

The control system is *repetitive* and functions by means of *action* (X acts on Y) and *reaction* ($E(Y)$ acts on X); through a certain number of interventions (iterations) on the *control lever* the control system tries to achieve an objective (*goal-seeking systems*) or respect a constraint or limit (*constraint-keeping systems*).

By observing the model in Fig. 2.2 according to systems thinking logic, it is clear that the control system is a *Ring*; that is, it produces a balancing loop through a logical structure of the model containing five fundamental variables and three functions.

The five variables of the model in Fig. 2.2 are as follows[1]:

1. The *variable to control* "Y_t," or *measured output*, which represents the *governed* variable (the *passive* variable) which characterizes the behavior of the entire system as it drives this variable to achieve the objective.

---

[1] In the figures in this book I shall use the symbols $\Delta(Y)$ or $E(Y)$ indifferently. Moreover, in order not to complicate the use of symbols, when it is understood that we are talking about dynamic variables (over time) I shall write X, Y, E and D, without indicating the variable "t" as a subscript.
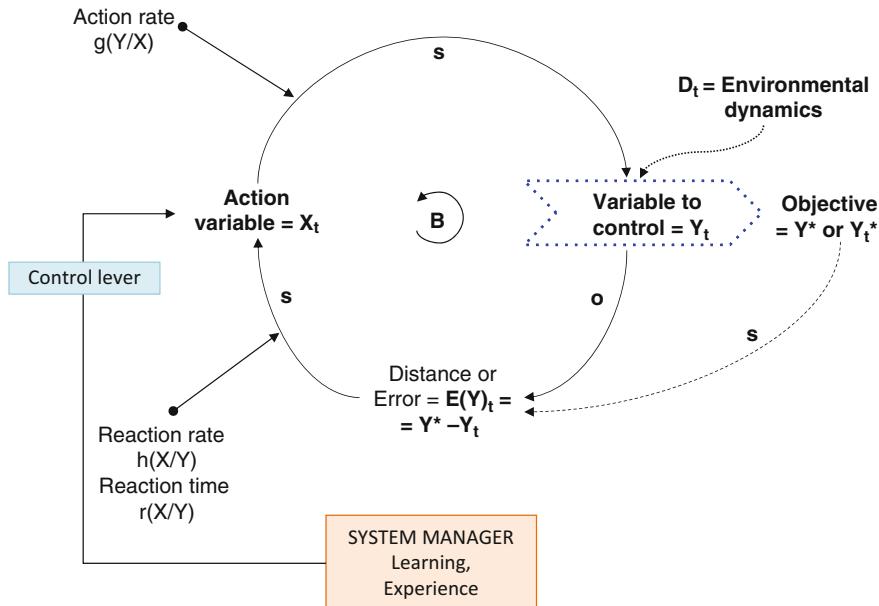
Fig. 2.2   Standard model of a one-lever control system

2. The *objective*, or *set reference*, which may be a time variable "Y_t^*" (tracking objective) or, as a particular case, a constant value "Y^*" (goal or constraint) that must be reached and possibly maintained by the variable to control; without specifications, I shall consider the objective as a constant value over time: "Y^*."
3. The error (deviation*,* gap, variance), $E(Y)_t = Y^*_t - Y_t$, which represents the *distance* between the values of the objective Y^* and those of Y_t; the causal links between Y and $E(Y)$ have direction "**o**" since, if we increase the value of Y, the error with respect to Y^* decreases; the causal links between Y^* and $E(Y)$ have direction "**s**" since, if we increase the level of the objective, we also increase the distance from Y.
4. The *action variable* or the *control lever* "X_t" (*governing* or *active* variable) which must be operated to modify "Y_t"; the causal link between X and Y has direction "**s**" since we assume that, in order to increase Y, we must increase X (the opposite relation with be considered below).
5. The variable "D_t," which indicates the possible external Disturbances connected to environmental Dynamics ("D" could also be a disturbance constant which acts one or more times); Fig. 2.2 does not indicate any direction of variation of the causal link between D and Y, since this direction can be either "**s**" or "**o**" based on the nature of D.

   Taking into account the direction of variation between the constituent variables, I shall define the model in Fig. 2.2 as a control system with structure [**s-o-s**].

The three functions needed to determine the dynamics of the system are the following:

1. The *action function* "$g(Y/X)$" (gain) indicates the measurement of the variation ΔY_t for each variation ΔX_t; if $g(Y/X) = g$, a constant, then the action function becomes simply an *action rate* which indicates the variation in Y for each unit of X; the accuracy of the *action function* or *rate* is fundamental for the efficiency of the control system and depends on the instruments available in relation to the type of system and on the control objectives.

2. The *reaction function* "$h(X/Y)$," or *reaction rate* "h," indicates the variation in X for each variation in E, and thus the variation in Y; usually "$h = 1/g$"; in this case the system is perfectly *symmetrical* to the control; in general, however, the system can be controlled even if "$h \neq 1/g$"; in general the lower "h" is with respect to "$1/g$" the easier is the control, since the variations of X as a function of $\Delta(Y)$ are attenuated, thereby producing more dampened impulses for X; if "$h > 1/g$," the system can produce increasingly greater oscillations. Section 2.16.11 considers the case of asymmetrical control systems.

3. The *reaction time* "$r(X/Y)$," or the *time to eliminate the error* (or the percentage of error which is eliminated in each step); this function indicates the speed with which the control system moves toward the objective; if $r(X/Y) = r$, a constant, a reaction time of $r = 1$ indicates an *immediate* but *sudden* control; a reaction time $r = 2$, 3, … makes the control *less fast* but *smoother*. $r = 2$, means that in each step a percentage of $(100/r)$ is eliminated. In order not to produce sudden adjustments, the variations of the error should be gradual and $r > 1$. The *reaction time* plays a fundamental role in control systems that produce their effects on people; the manager must therefore balance the need for rapid control (if a pedestrian is crossing the street a sudden breaking is required) with that of graduality (if you are giving a ride to your old granny it is better to avoid sudden breaking).

Notice that the assumptions that $g(Y/X)$ and $h(X/Y)$ are *constant coefficients* (that is, *action rates*), for example, $g = 4$ and $h = 0.25$, do not represent a serious limitation for control systems; quite the opposite: the presence of constant coefficients are found in a large variety of control systems with linear dynamics. In many other cases, however, $g(Y/X)$ and $h(X/Y)$ are functions of X_t and/or Y_t and/or E_t, so that the system can also have nonlinear dynamics, as occurs when, for example, $g(Y/X) = 1.5 - 0.1X_{t-1}$; in this case the dynamics of Y_t would depend not only on the constant *action rate* ($g = 1.5$) but also on the new variable coefficient "$0.1X_{t-1}$."

What matters most is to understand that the dynamics of the control lever—and as a consequence of the entire system—depends not so much on the value of Y_t as on the "distance" between Y^* and Y_t. The concept of "distance," or "Error," is fundamental in all feedback control systems; in fact, Norbert Wiener, the founder of cybernetics, proposes this concept as the basic mechanism of such a form of control:

> Now, suppose that I pick up a lead pencil. To do this, I have to move certain muscles. However, for all of us but a few expert anatomists, we do not know what these muscles are; and even among the anatomists, there are few, if any, who can perform the act by a conscious willing in succession of the contradiction of each muscle concerned. On the

contrary, what we will is to pick the pencil up. Once we have determined on this, our motion proceeds in such a way that we may say roughly that the amount by which the pencil is not yet picked up is decreased at each stage. This part of the action is not in full consciousness. To perform an action in such a manner, there must be a report to the nervous system, conscious or unconscious, of the amount by which we have failed to pick up the pencil at each instant (Wiener 1961, p. 7).

Systems thinking has made the concept of distance operative simply by creating the new variable $E(Y)_t = Y^* - Y_t$, which tells us "by how much we have exceeded Y^*" or "how close we are to Y^*" (Fig. 2.1).

Let us agree to denote systems that function according to the model in Fig. 2.2 also as *Decision and control systems*, since the control of Y_t implies that explicit decisions regarding changes in X_t are made by the system manager. A control system is defined as *automatic*, or *cybernetic*, if the entire system can be structured and programmed so that its functioning is independent of human decisions and interventions; the *cybernetic* system appears to be closed within itself, and the only link with the environment is the specification of the objective Y^* by the *manager*. We shall return to these aspects of control systems in Chap. 3. We should note at this point that control systems operate automatically for the most part, at microscopic and macroscopic levels; for this reason they are not easily observed or understood.

2.3   The *Ring* in Action: The Heuristic Model of a Control System

On the basis of this minimal set of variables and functions, represented by the logical model in Fig. 2.2, we can derive the simple *heuristic model* that describes the recursive equations producing the behavior of the system (assumed to be without delays) through the dynamics of its variables.

For simplicity's sake, we assume that (a) the time scale is represented by $t = 0$, 1, 2, ... n, ...; (b) $g(Y/X) = g$ is constant; (c) the system is symmetrical and $h = 1/g$; and (d) $r(X/Y) = r$: that is, $r = 1$ or $r = 2$, etc.

- At time ($t = 0$), detect the *initial state*, Y_0, of the variable to control, and the objective, Y^*.
- At time ($t = 0$), quantify the "distance," or error, between the initial state and the objective to achieve, calculating the initial deviation:

$$E(Y)_0 = Y^* - Y_0 = \Delta(Y)_0 \tag{2.0}$$

- At time ($t = 0$), determine the value X_1 to be assigned to the action variable in the subsequent moment $t = 1$; this value can be assigned by the manager of the system at his discretion, based on his experience, or it can be determined more efficiently by multiplying the deviation by the *reaction time* "*h*" (assumed to be constant):

$$X_1 = X_0 + E(Y)_0\, h \tag{2.1}$$

- At time ($t = 1$), determine, using the *action rate* "g" (assumed constant), the first value Y_1 for the variable to control:

$$Y_1 = X_1 g + D_1 \qquad (2.2)$$

- Again at time ($t = 1$), determine the new error:

$$E(Y)_1 = Y^* - Y_1 \qquad (2.3)$$

  where $E(Y)_1 < E(Y)_0$
- At time ($t = 2$) the action variable is adjusted by adding the correction factor, calculated using the *reaction rate* "h," according to the equation:

$$X_2 = X_1 + E(Y)_1 \, h \qquad (2.4)$$

- As there are *no delays*, again at ($t = 2$) calculate the new value for the variable to control, using the equation:

$$Y_2 = X_2 g + D_2 \qquad (2.5)$$

- Proceeding recursively, at discrete moments ($t = n$) and ($t = n + 1$) we calculate the new values for Y and X from the equations:

$$Y_n = X_n \, g + D_n \qquad (2.6)$$

$$E(Y)_n = Y^* - Y_n \qquad (2.7)$$

$$X_{n+1} = X_n + E(Y)_n \, h \qquad (2.8)$$

- Having assumed the values of the reaction time are $r = 1$ or $r = 2$, etc., we can modify (2.8) (also modifying the preceding relations) to get the following relation:

$$X_{n+1} = X_n + \frac{(Y^* - Y_n)h}{r} \qquad (2.9)$$

In this sense "r" signifies *time required to eliminate error*. Note that since (2.9) allows us to adjust the variables with respect to the errors of the preceding instant by multiplying the error by ($1/r$), the system in theory will never achieve equilibrium. It is thus necessary to define a *tolerance* which, when reached, will lead the system to automatically ignore the residual error and consider the objective value to have been reached.

The system achieves a *stable equilibrium* when the deviation $E(Y)_{n*} = (Y^* - Y_{n*})$ is zero for all moments subsequent to "n^*" and the variable Y_{n*} reaches and maintains the value Y^*; or when there is a deviation opposite in sign that the system eliminates by acting appropriately with new values for X. This stability can be *long-lasting* but, due to the action of the *law of dynamic instability* that characterizes all systems from the system thinking perspective (see Sect. 1.6.3), this cannot be indefinitely maintained,
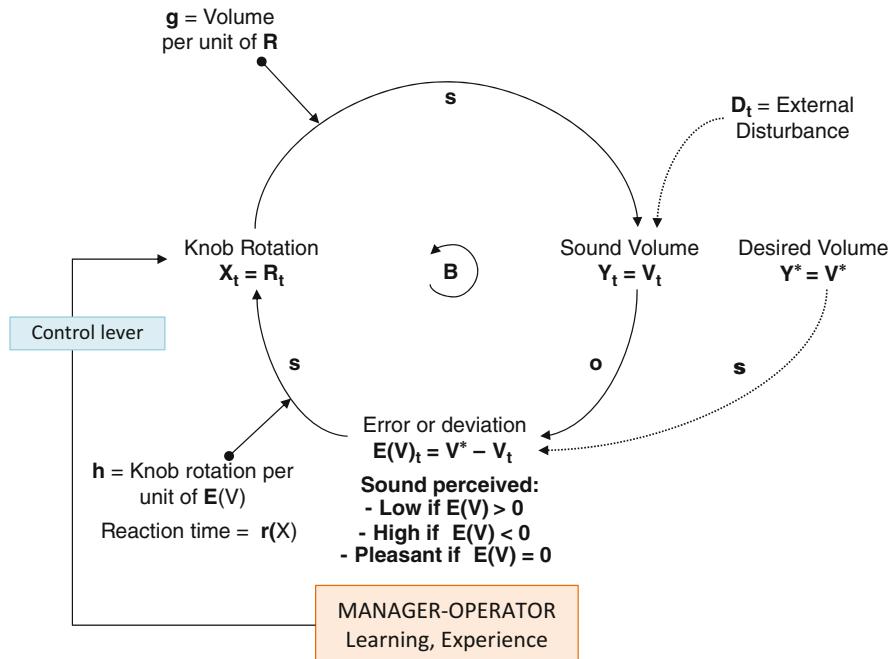
Fig. 2.3 Audio control system with one continuous lever without delay

since normally some disturbance, D_n, occurs that requires new adjustments. The above elements are usually sufficient by themselves to describe the general model of system control without delays, using the symbols of systems thinking.

Regarding the way the values of the variables X_n and Y_n are calculated, I would note that, through (2.6) and (2.8), the cumulative values of variables X_n and Y_n are always determined. One alternative is possible: calculating the cumulative values for X_n or for Y_n alone. An equivalent alternative is possible: calculating the difference values: $\Delta X_n = X_{n+1} - X_n$ and $\Delta Y_n = g\Delta X_n$.

Naturally, in both cases the dynamics always leads to equilibrium, though the way in which it is represented changes; for example, in the numerical tables for the audio and shower temperature models shown in Figs. 2.3 and 2.9 below, in order to facilitate the understanding of the dynamics I have chosen to represent the dynamics of Y_n and only the variations in X_n for each "n."

2.4 Some Technical Notes

Even if the control system model has been completely outlined, some technical notes are useful to clarify or supplement several elements of the model in Fig. 2.2.

First technical note. A control system is structured to *achieve* the value Y* through a succession of "adjustments" to X; therefore, there can be no control system if we

do not have some type of objective. Whatever the nature of the objective, it can be made explicit or implicit. Y^* represents an *implicit* objective if it entails a constraint or limit which, though not defined, is the result of the "nature of things." Our normal body temperature is an implicit objective (constraint) of our physiological system, which maintains the temperature over time by acting on "sweating" and "chilling." The temperature of 22 °C we want to maintain in our office with a heater is an *explicit* objective. Thus, we can define a control system in simple terms as a *Ring* that operates repetitively to allow the Y_t to achieve the *explicit* or *implicit* objective independently of its significance.

Second technical note. I observe that in very large systems composed of many subsystems—whose CLD is made up of a large number of variables—it is not always easy to recognize the existence of control systems. There is, however, a simple *rule*: a control system exists whenever we can recognize in the general system we are observing and modeling the presence of a variable that takes on the significance of an error with respect to a value Y^*, appropriately defined, that serves as an implicit or explicit objective.

Third technical note. control systems are dynamic systems that achieve or maintain the objective in a *succession of iterations*. Nevertheless, there are also control systems for which the objective must be achieved in a *single* cycle, as for example in hunting, an activity where normally the hunter has only one shot to catch the prey (objective), or in industrial or artisan activities, where a given worker must carry out an activity that leads to irreversible effects. Also falling within this category are all those systems where a decision must be made that produces a single, non-modifiable effect. I shall call these control systems *one-shot systems* or simply *decisional systems*. Decisional processes, in the traditional sense of the term—choosing the best alternative for achieving an objective—can thus be considered as special cases of control processes (see Sect. 5.8). Henceforth, unless otherwise specified, the term control system will mean a *pluricycle system of decision-making and control*: a *Ring* that rotates several times.

Fourth technical note. The control system in Fig. 2.2 has a simple feature: the action on X_t is proportionate to the error (in the various ways that will be examined below). These systems can be defined as *proportional control systems* in that they correct X_t, and thus Y_t, in a measure that is proportionate to $E(Y_t)$. In some cases it can be useful to quantify the action on X_t not so much in terms of the error but of the variation in $E(Y_t)$; the corrective measure is proportionate to the rate of variation more so than to the amount of the error. Systems that adopt this form of control can be called *derivative control systems*; they permit a more "precise" regulation of the variable Y_t: the smaller is the reduction in $E(Y)_t$, the more dampened is the variation in X_t (for example, approaching the dock the speed of the boat's motor is reduced). There is also a third possibility: quantifying the reaction on X_t not based on the error dynamics, at every instant, but on the dynamics of the accumulated error, from the beginning of the control and continuing on through each successive instant. For this reason these systems are known as *integral control systems*, and they can be applied

to phenomena for which the error indicates an overall magnitude to be "bridged" (volume, distance, etc.); in this case the intervention on X must be proportionate to the magnitude that is "narrowed." In many cases it is also necessary to maneuver the control lever to take into account possible future errors, in order to bring forward the control and make it more efficient. A control system that seeks to control Y by setting values for the X lever so as to simultaneously take into account three measures: (1) the *size* of the error at every instant (P, *proportional* to the present error), (2) the total (integral) of the past errors (I, integral of past errors), (3) the prediction of future errors (D, *derivative* of the error; that is, the slope of the error over time) is called a proportional-integral-derivative control system or PID control system (Sontag 1998; Zhong 2006).

Fifth technical note. The terms control and regulation are often used with various meanings. For example, in *control theory* (Ogata 2002; Sontag 1998, p. 372) the term *regulation* indicates the form of control typical of systems that must maintain the variable Y stable *around* the values of an objective (*set point*) which is set once and only varied occasionally by the system's *manager*. Along with *systems of regulation* do we have *position control systems* (or servomechanisms), whose function is to modify the variable Y through X in order to reach an objective than can also vary over time (I shall deal with these systems in Sect. 5.4). A second meaning, in addition to the two just mentioned, is based on the distinction between the control process and the technical system ("machine") that carries this out, in the sense that: "I cannot control the temperature because the mixer faucet is not well regulated." Here the *control* regards the directional process of Y_t in moving toward the objective; the *regulation*, on the other hand, concerns the parameters for the functioning of the "machine" that carries out the *control*. It is clear that these two concepts are linked, in that the process depends on the "machine." *Ergo*: the *control* depends on the *regulation*. In order to avoid ambiguity, I shall use the term *calibration* rather than *regulation* to refer to the specific techniques of the "machine" that carries out the control process. According to a third meaning, control implies the attainment of an objective set by man; regulation instead represents a "natural" control, without human intervention, as in the case of a mutual regulation of populations (Sect. 7.2). In this book, to avoid ambiguity I shall use the term "control" with the more general meaning of indicating the entire *logical* process that leads Y to achieve the objective Y^*, even assuming deliberate or systematic variations in the latter. I shall use the term "regulation" to indicate the choice of values for X which are necessary to modify Y in order to eliminate the error, as in the expressions "to control the temperature (variable Y), regulate the mixer faucet lever (control lever for X)"; "regulate the speed (control lever for X); otherwise we'll go off the road (variable Y)"; "get going on the rudder (control lever for X) and change direction (variable Y); otherwise we won't enter the harbor (objective)." I shall use the term "calibration" to indicate the control on the parameters of the "machine" that implements the control process.

Sixth technical note. The logic of a *closed-loop* or *feedback control* described so far is fundamental for the arguments in this book. Nevertheless, we can also design

control systems defined as *open-loop*, *feedforward*, or even *commands* where the value of the control lever is determined only at $t = 0$ based on a *model known to the manager* by which he regulates the causal relations between X and Y, without further interventions on the trajectory of Y, in order to eliminate any error (Sect. 2.16.9 will deal with this topic further).

The next sections present several models typical of control systems and describe simple systems we have all encountered in our daily lives, in order that the reader can verify the *modus operandi* of those systems and understand how they perfectly achieve the control process.

2.5 Continuous Single-Lever Control System Without Delays

It is useful to begin by presenting the simplest control system, with a single control lever, X_t, with continuous variation within a preestablished range of values. This system regulates a variable, Y_t, which is also characterized by continuous variation and which reacts instantly—that is, without delay—to the variations of the control lever in order to achieve an objective, X^*, fixed over time and with a constant and known action rate, g, that regulates the variation of Y_t, assuming the reaction time is $r = 1$.

The continuity of the variables, the absence of delays, a fixed objective, and a fixed action rate are all features which make the control very simple. Knowing "g" and observing the amount of $E(Y)$, the manager assigns values to X_t in successive instants ($t = 1, 2, \ldots$) which allow Y_t to reduce the distance from Y^* and to eliminate $E(Y)$ at a certain instant t^*. This control system can be easily simulated, and if the manager had access to a precise error *detector*, while knowing $E(Y)_0$ and the action rate "g," he would be able to calculate the new value X_1 to the extent necessary to obtain the perfect control already by instant one, since the reaction time is $r = 1$. The error elimination would be immediate, and thus very abrupt. To make the control less abrupt we could introduce a reaction time of $r > 1$.

Though it has a quite elementary logical structure, this type of system is very common, as our daily experience will confirm. We find such a control system when we have to adjust the volume on our radio, TV, stereo, cell phone, iPod, etc., using a knob with a continuous rotation. We observe the same logic when we have to change the direction of our car by turning the steering wheel or when we wish to change the water flow by turning the faucet, and so on.

When we listen to the radio we can judge the volume level as being pleasant, too high or too low; thus, there is an optimal, subjective level V^* which represents the objective of the control. Given the physical structure of the "machine" producing it—which we assume to be a black box—we regulate the sound volume by rotating the regulating knob by a certain number of "degrees" or "notches" displayed on a graduated scale, as shown in Fig. 2.3.

The knob rotations, R_t, represent the action variable; that is, the *control lever*. The sound volume, V_t, is the variable to control in order to achieve the objective V^*.

A clockwise rotation (cause) produces the effect of "increasing the sound volume"; a counterclockwise rotation "reduces the volume." If we define a clockwise rotation as positive (+) and a counterclockwise one as negative (−), then a causal relation in direction "**s**" exists between R and V.

The sound control is achieved through a succession of decisions that transmits to R the impulses $\Delta^+R_t =$ "turn right" or $\Delta^-R_t =$ "turn left," which depend on the sign and perceived entity of the error, $E(V)_t$, of the actual volume compared to the desired one. Once the objective (V^*) of a pleasant sound is reached, the system remains stable, in the sense that no further rotation ΔR_t is applied to the knob, unless the user modifies the objective V^* or some *external disturbance* to the sound volume makes a correction necessary. If the control occurs *manually*, "by ear," a normal manager–operator would be able to achieve the objective within a few time units, though with a certain tolerance. If we assume that the control system is *automatic* and that the manager–operator only sets the objective V^*, then the dynamics of the variables that make up the logical structure of the control system for the volume can be derived from (2.0) to (2.8).

Let us assume the following initial inputs:

- $R_0 = V_0 = 0$; the initial values of the control lever and the volume are zero.
- $V^* = 40$ dB.
- $g = \frac{1}{2}$; if we assume the knob rotates only a maximum of $120°$ and that the maximum attainable volume is 60 dB, then the action rate $g = \frac{1}{2}$ means we have to rotate the knob $2°$ to obtain a volume increase of 1 dB.
- $h = 2 = 1/g$; this means that each unit of error $E(V)$ requires a correction of $2°$ to the knob.
- Reaction time $r = 1$.

These values allow us to produce the following iterations that carry out a control process in a single step:

- At time $(t = 0)$, quantify the "distance" between the initial state and the objective to achieve, calculating the initial deviation:

$$E(V)_0 = Y^* − Y_1 = 40 − 0 = 40 \;\; \text{decibels}$$

- At time $(t = 0)$, determine the value R_1 to be assigned to the control lever at $t = 1$; this value can be assigned by the manager–operator of the system at his discretion, based on his experience, or it can be determined more efficiently, and automatically, by multiplying the error $E(V)_0$ by the *reaction rate*, $h = 2$:

$$R_1 = R_0 + E(V)_0\, h = 0 + (40 \times 2) = 80 \;\; \text{degrees of rotation of the knob}$$

- At time $(t = 1)$, determine, using the *action rate* $g = \frac{1}{2}$, the first value V_1 for the sound volume:

$$V_1 = R_1 g + D_1 = (80 \times \tfrac{1}{2}) + 0 = 40 \;\; \text{decibels}$$

| g = decibels/dial | | | 0.5 | $V_0 =$ | 0 | r(R) = | 2 |
| h = 1/g = dial/decibels | | | 2 | $V^*_0 =$ | 40 | tolerance | 0 |

t	delta R_t = rotation	X_t = cumulative rotation	V_t = volume/ decibels	V^*_t = objective	variation in objective	D = disturbance	DEVIATION $E(V)_t = V^*-V_t$
0	0.00	0.0	0.0	40.0	0	0	40.0
1	40.00	40.0	20.0	40.0	0	0	20.0
2	20.00	60.0	30.0	40.0	0	0	10.0
3	10.00	70.0	35.0	40.0	0	0	5.0
4	5.00	75.0	37.5	40.0	0	0	2.5
5	2.50	77.5	38.8	40.0	0	0	1.3
6	1.25	78.8	39.4	40.0	0	0	0.6
7	0.63	79.4	39.7	40.0	0	0	0.3
8	0.31	79.7	39.8	40.0	0	0	0.2
9	0.16	79.8	39.9	40.0	0	0	0.1
10	0.08	79.9	40.0	40.0	0	0	0.0
11	0.04	80.0	40.0	40.0	0	0	0.0
12	0.02	80.0	30.0	50.0	10	-10	20.0
13	20.01	100.0	40.0	50.0	0	-10	10.0
14	10.00	110.0	45.0	50.0	0	-10	5.0
15					0	-10	2.5
16					-10	0	-18.7
17					0	0	-9.4
18					0	0	-4.7
19					0	0	-2.3
20					0	0	-1.2
21					0	0	-0.6
22					0	0	-0.3
23					0	0	-0.1
24					0	0	-0.1
25					0	0	0.0

Chart (overlapping rows 15–25): **AUDIO VOLUME CONTROL CONTINUOUS** — y-axis: 70.0, 60.0, 50.0, 40.0, 30.0, 20.0, 10.0, 0.0; labels: **Objective**, **Volume/decibel**; x-axis: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

Fig. 2.4 Dynamics of the variables in the volume control with a continuous lever

- At time ($t = 1$) determine the new deviation:

$$E(V)_1 = V^* - V_1 = 40 - 40 = 0 (\text{control achieved}).$$

Observe that the automatic control of the sound volume has led to the immediate attainment of the objective at instant $t^* = 1$. If we desire a more gradual increase in volume we could construct the control system using a reaction time of $r(R) = 2$. Figure 2.4 shows the numerical progression of the dynamics of the variables under this hypothesis. The "control panel" at the top of the table indicates the parameters the manager–operator must specify, including the volume objective and the action rate "g." For convenience sake, the error is shown in the last column. It is important to realize at this point that the discomfort we experience when the volume is too high or too low is the *signal* that there is a *gap* with respect to the level we judge to be pleasant; this sense of discomfort is a *symptom* that creates the need to control V; thus it is subjective, often depending on environmental circumstances as well.

As shown in the system's control panel in Fig. 2.4 (first two lines at the top) we have assumed that the *reaction time* is $r(R) = 2$ time units; this means that rotating the dial $80°$ cannot be done in a single action, since the error reduction must occur over 2 time units. As a result, and as we can read from the line corresponding to $t = 1$, the rotation is not $R_1 = 80°$ but $R_1 = 40$, which produces an increase of only $V_1 = 20$ dB of sound (third column, $t = 1$), thereby reducing $E(V)_1$ by half. At $t = 2$, a $40°$ rotation of the dial is to occur in order to eliminate the error. Since $r(R) = 2$, the dial rotation is only $R_2 = 20°$, a volume increase of only 10 dB, which makes the perceived volume $V_2 = 30$ dB (third column, $t = 2$); the system continues on in this way until instant $t* = 10$, when the error is eliminated. There is thus a nonlinear control, as we can easily see from the graph in Fig. 2.4.

By observing the values from column 2 ($X_t =$ "cumulative rotation") we note that the sum of the partial rotations from $t = 1$ to $t = 10$ is precisely $X_{10} = 80°$ and that corresponding to this value of the action lever is a volume level of $V_{10} = 40$ dB. At $t = 10$ the objective has been achieved and the error eliminated. If the reaction time had been $r(R) = 4$, the adjustment would have occurred for one-fourth of $E(V)$ for each period.

The concept of *reaction time* may seem unrealistic since we are not used to thinking that a control system does not immediately adjust X to eliminate the error and achieve the objective $Y*$. However, the opposite is true. Experience shows that normally the corrections made by the control system are not violent but occur gradually over several time instants. If you are not convinced of this, consider how irritating it would be if a friend of yours turned on the radio and the radio suddenly rose from zero to 60 dB, making those present jump. And consider how annoying, or even harmful, it would be if, by turning the shower mixer, the water immediately became hot or cold; you would jump out of the shower. Or if, when you get into your friend's car, it immediately reaches cruising speed, slamming your body against the seat. However, with single-lever control systems, especially those that gradually produce their effects on people, the adjustments are normally fairly gradual, and this gradualism depends on the non-instantaneous *reaction time*, which indicates by how much the error must be reduced in each interval.

Returning to the model in Fig. 2.4, the penultimate column shows that from $t = 12$ to $t = 15$ there is an external disturbance (for example, the nearby church bells are ringing) which reduces the perceived volume by 10 dB. I have therefore assumed that at $t = 12$ the manager–operator decides to increase the objective from $Y* = 40$ to $Y* = 50$ dB in order to counter the disturbance effect, the volume V rapidly increases as the graph shown in Fig. 2.4. Once the disturbance ceases the audio system produces a decibel level that is so annoying that the manager–operator decides to once again reduce the objective to $V* = 40$ dB. The control system gradually reduces the dial rotation, returning the decibel value to that of the objective, which is reached at $t = 24$, with a minimal tolerance.

The continuous-lever control system described in the model in Fig. 2.3 is, despite its simplicity, entirely general and capable of representing numerous regulation processes. I shall here mention a few of these processes we all have experienced.

System name	Y = variable to control	X = continuous control lever
Stereo system	Volume level	Dial or regulator
Lamp	Level of light	Electricity regulator
Roller blind	Amount of light that enters	Lift cord
Eye	Amount of light that enters	Pupil size
Automobile	Direction of movement	Steering wheel rotation
Automobile	Speed (increase)	Pressure on gas pedal
Sailboat	Direction of movement	Rudder (angle of rudder)
Sailing ship	Direction of movement	Steering wheel (rotation)
Military drone	Direction of movement	Rotation of joystick
Nuclear engine	Power produced	Number of uranium bars
Automobile	Arrival time	Average speed

2.6   Discrete Single-Lever Control System Without Delay

Other control dynamics are possible which depend on the type of *physical apparatus* used in the regulation process. In fact, along with *continuous-lever* systems there are *discrete-lever* control systems, which are equally simple and widespread.

Figure 2.5 presents a variant of the audio control in which the control lever is not a continuous rotating dial but the button on a remote control device (that can be pressed in two directions for several successive instants for a certain period) to increase or decrease the audio by a *constant amount* so that, at the same time, the error $E(V)_t$ falls at each instant by the same discrete amount at each subsequent pressing of the button. The control lever with discrete variations which, at each variation, varies the control variable by a discrete quantity (the movement of the lever) represents a "fixed-turning" lever. When there is a discrete and fixed-turning lever, the reaction time is not set, since the speed of reaction is directly established by the manager–operator. As shown in Fig. 2.5, the dynamics of V_t, while linear, gives similar results to those in Fig. 2.4.

We can consider as *discrete* or *fixed turning lever* systems all those where the lever causes a variation in Y_t proportional to the action time, so that each action time unit of the lever is entirely equivalent to the pressure of a regulator button. If a tap produces a constant water flow so that the water volume is, for example, $g = 1$ l per second, then I can fill a container with a volume of $V^* = 10$ l by keeping the lever open (that is, the constant opening of the tap) for a period $T = 10$ s, which would be equivalent to producing the water volume for 1 s ten times. The similarity to the discrete button-lever is totally evident.

We can include in this class all control systems whose lever is a constant flow per unit of time, $f_t = f =$ "constant flow/time unit," where $t = 0, 1, 2, \ldots$ (1 l per second, 1 m^3 per day, 100 m per minute, etc.) and the variable to control is a flow integral in a defined interval T, $Y_t = \int f \mathrm{d}t = (f \times T)$: liters accumulated in 20 s; gas consumed in 40 days; distance covered in 30 min, etc. Given these characteristics, the variations of the control lever can, in fact, correspond to the instants

g = fixed turning of the dial		V_0 =	0			
	10	V*_0 =	40			

t	$R =$ decibels/click	$V_n =$ volume/decibels	$V^*_n =$ objective	variation in objective	$D_n =$ disturbance	DEVIATION $E(V)_n = V^*_n - V_n$
0	0.00	0.0	40.0	0	0	40.0
1	10.00	10.0	40.0	0	0	30.0
2	10.00	20.0	40.0	0	0	20.0
3	10.00	30.0	40.0	0	0	10.0
4	10.00	40.0	40.0	0	0	0.0
5	0.00	40.0	40.0	0	0	0.0
6	0.00	40.0	40.0	0	0	0.0
7	0.00	40.0	40.0	0	0	0.0
8	0.00	40.0	40.0	0	0	0.0
9	0.00	40.0	40.0	0	0	0.0
10	0.00	40.0	40.0	0	0	0.0
11	0.00	40.0	40.0	0	0	0.0
12	0.00	30.0	50.0	10	-10	20.0
				0	-10	20.0
				0	-10	20.0
				0	-10	20.0
				-10	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0
				0	0	0.0

AUDIO VOLUME CONTROL (discrete) — Volume/decibel

Fig. 2.5 Sound volume control with discrete control lever

$X_t = t$, which refer to the use of the lever; the higher X_t is (the number of instants of flow emission), the greater Y_t becomes (overall flow volume).

Even the system in Fig. 2.5 follows this rule. In fact, we can observe that the discrete lever was pressed four times, or four instants, and that V_t takes on the value $V_t = 4 \times 10 = 40$. This interpretation allows us to broaden considerably the number of *discrete* systems with a *fixed-turning lever* that we can encounter in the real world, which can be represented by the model in Fig. 2.5. Some of these systems are indicated in the following table.

System name	Y = variable to control	X = discrete control lever
Stereo system	Sound level	Pressing the remote control
Lamp	Level of light	Variator with relay
Sailboat	Direction of movement	Rudder angle
Cylindrical cistern	Amount/depth of liquid	Flow emission time
Automobile	Distance travelled	Hours at constant average speed

2.7 Control System with On–Off Lever

We can further widen the class of systems that clearly follow the *modus operandi* described in Sect. 2.6 by including as well the type of control system defined as *on–off*. In fact, for such systems we can assume that the action of the lever corresponds to a duration T formed by a succession of consecutive instants, $X_t = t$, in each of which the variable Y_t undergoes a variation (continuous or fixed-turning) until the objective is achieved. The control lever is activated (*on*) when the error occurs and is deactivated (*off*) when the error is eliminated. If an external disturbance should produce a new error, the lever will once again activate (*on*) until the new error reaches zero (*off*). Thus, an external observer would perceive the activity of these systems as a succession of activations and deactivations of the control process; that is, a succession of periods T where the lever is operating, in each of which there is a variation in Y_t which is necessary to eliminate the error, following a typical *on–off* behavior, which accounts for the term *on–off* systems.

Due to the simplicity of their logic, *on–off* systems are quite common. Some of those which are immediately recognizable in our home environments are presented in the following table.

System name	Y = variable to control	X = continuous control lever
Air conditioner	Environmental temperature	On–off compressor
Refrigerator	Temperature below zero	On–off compressor
Radiator fan	Heat of radiator	On–off fan motor
Forehead sweating	Body temperature	On–off sweat gland
Chills	Body coldness	On–off muscular vibration
Windshield wipers	Transparency of windshield	On–off wiper oscillator
Brightness regul.	Brightness monitor	On–off led power
Fireplace	Room temperature	I/O feeding fire
Fan	Air temperature	I/O fan power supply

On–off systems usually function *automatically*, since the variable to regulate undergoes variations from external disturbances whose intensity, duration and starting time cannot be easily predicted. For this reason, with *on–off* systems—for example, air conditioners—the activation and deactivation of X_t does not depend on decisions by the manager but on the magnitude of $E(Y)_t$ and the control systems operating program.
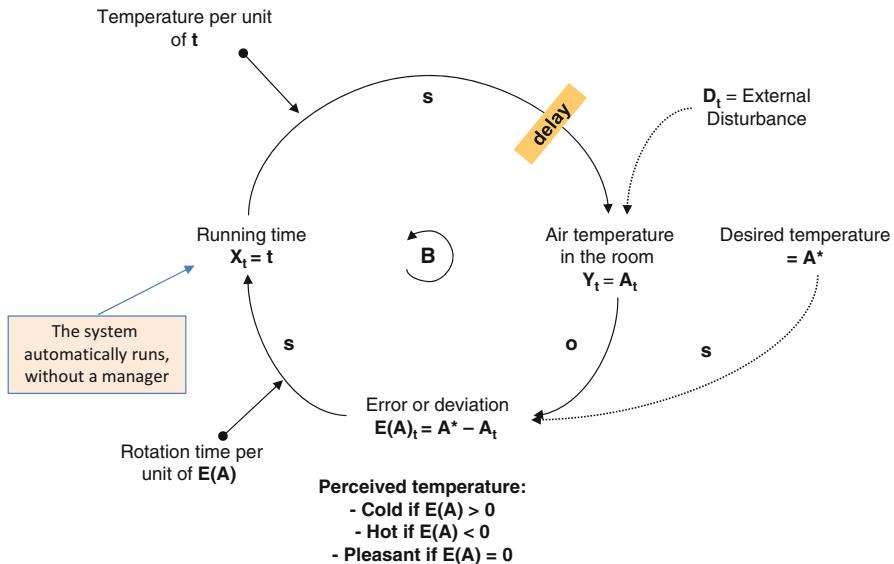
Fig. 2.6 Model of an automatic on–off air conditioner

Figure 2.6 presents an initial standard model to examine of an automatic *on–off* control system for regulating room temperature by means of a basic single-lever air conditioner.

We immediately observe that this model is entirely similar to that in Fig. 2.5, except for the fact that the *reaction rate*, *h*, corresponds to the command to turn on the conditioner, which functions automatically for a number of instants $X_t = t$, in each of which cold air is emitted at the temperature indicated by the *action rate*, *g*; the actual set-point temperature is gradually lowered to the objective *A**. The desired temperature is memorized through an appropriate graduated dial (or a digital display) indicating the various temperatures which, when set at a fixed point, memorize the objective *Y**.

If outside disturbances should once again raise the temperature, the regulator would again turn on the compressor for a sufficient interval to reach the objective. It is clear that over a sufficiently long period the control system produces a succession of compressor activations (on) and deactivations (off). Obviously when producing an air conditioner, as well as other *on–off* systems, account is taken of a tolerance interval for Y_t with respect to the objective, so that the compressor does not have to turn on and off too frequently.

On–off systems can also operate with a *delay*, though the delayed start is usually compensated for by an equally delayed turning off. The *reaction time* also appears not to be significant since its effect is compensated for, once again by the length of the activation period.

A non-automatic *on–off* system—that is, one where the activations and deactivations of X_t (which determine *Y*) depend directly on the manager's decisions—is

often indicated by the symbol I/O: "I" for *on*, "O" for *off*. We can find such non-automatic systems in, for example, country homes where a room is heated by adding logs to the fireplace for a certain period or cooled down by turning a fan on or off; or in vintage cars without radiator fans, which must therefore stop from time to time to cool down the liquid in the radiator.

2.8 Continuous Single-Lever Control System with Delay

This section takes up control systems with a *delay* between the variation of X_t and that of Y_t. We shall construct a new standard model that allows us to introduce the concept of delay and to clarify how delays condition the behavior of control systems. Unless stated otherwise, it is assumed at all times that the variations in Y_t resulting from the variations in X_t are instantaneous; this means that the system takes no time to carry out its process or that the amount of time is negligible, exactly as in the example of turning off the radio dial, which seems to "instantaneously" change the volume, at most taking account of the gradualness resulting from the reaction time. When instead there is a *relatively long* and *unexpected* interval between the assignment of values to X_t, and the "answer" comes from $Y_{t+\Delta t}$, and not Y_t—with a non-negligible Δt—then the control system operates with a delay of Δt.

In practical terms familiar to all, we can state that "Δt" represents a delay when the interval is so long that the *manager* can decide to vary X *at least twice* before the *first* variation is produced in Y. Precisely for this reason, *delays* make it quite complicated for control systems to achieve their objective since, if we attribute two impulses, ΔX_1 and ΔX_2, to X in rapid succession, we can have an answer, $\Delta Y_{\Delta t}$, that is twice that of the desired one, which would then require a new impulse of opposite sign to X_t.

The presence of *delays* in models is indicated by the word "delay" appearing on the arrow as many times as there are impulses from X before the first response from Y, or by placing two short crosswise segments on the arrow.

It is important to realize that *delays* do not always depend on functional defects of the control system but, in most cases, are innate to the functioning of the "machines" that activate control processes (Sect. 2.9). Thus *delays cannot be eliminated*, arbitrarily reduced, or even ignored. Once cognizant of the existence of delays and the impossibility of eliminating them, *there is only one strategy open to us*: learning *to identify them* and, through experience, *reducing them in number* and *length*. Delays can also arise in the recursive equations in Sect. 2.3. A single *delay* means that the value of a variable to control, "Y_n," at a given instant, $t = n$, depends on the value X_{n-1} of a control lever rather than on X_n. The recursive equations (2.6)–(2.8) become:

$$Y_n = X_{n-1} \ g + D \qquad\qquad (2.6/1)$$
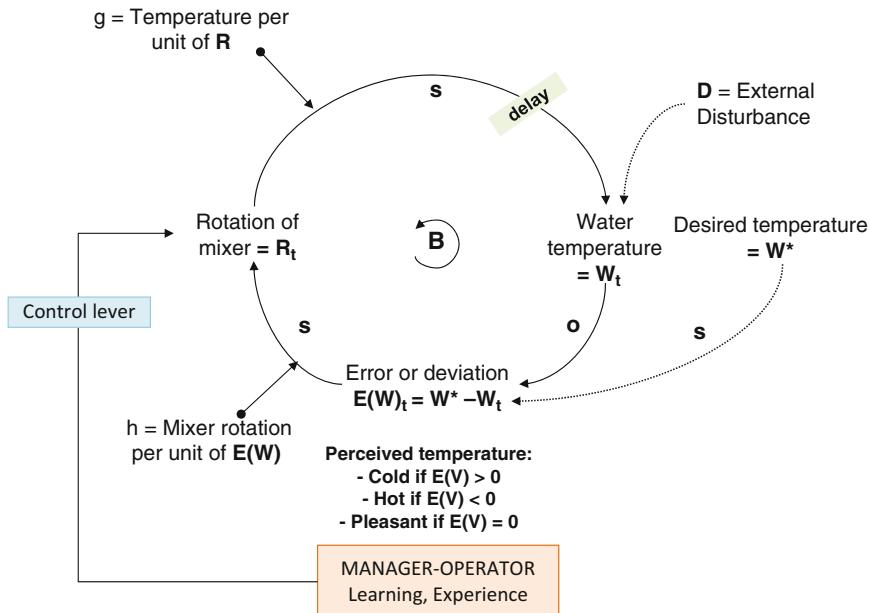
Fig. 2.7 Control system for regulating the shower temperature using a mixer

$$E(Y)_n = Y^* - Y_n \tag{2.7/1}$$

$$X_{n+1} = X_n + E(Y)_n \ h. \tag{2.8/1}$$

The *delays* mean that at $t = n$ the value of a variable to control, "Y_n," depends on the value X_{n-2} of a control lever rather than on X_n.

The recursive equations (2.6/1)–(2.8/1) become:

$$Y_n = X_{n-2} \ g + D_n \tag{2.6/2}$$

$$\Delta(Y)_n = Y^* - Y_n \tag{2.7/2}$$

$$X_{n+1} = X_n + \Delta(Y)_n \ h. \tag{2.8/2}$$

To understand the effect of *delays* we can consider, for example, the system that regulates the water temperature of a shower with a mixer (single-lever control), as shown in the model in Fig. 2.7, which represents a theoretical shower with a delay between the regulation of the mixer and the variation in the water temperature, exactly similar to the sound regulation model in Fig. 2.4.

Figure 2.8 presents a *numerical simulation*, using Excel, of the system in Fig. 2.7, assuming the "delay = 0." In Fig. 2.8 I have assumed that the shower temperature is regulated by a mixer that, for each degree of rotation, $(R = X)$,

g = temperature regulator

0,5 degrees/notches

Rotation at t_0

0

Rotation of

R_t = rotation in mm

B

learning
experience

s

2,00   per unit of E(W)

h = mixer rotation

s

NO delay

0

Temperature

W_t = temper. (C°)

o

Deviation
$E(W) = W^*-W$

External Disturbance

temperature at t_0

0

Desired
temperature = W*

28

s

2,00   reaction time
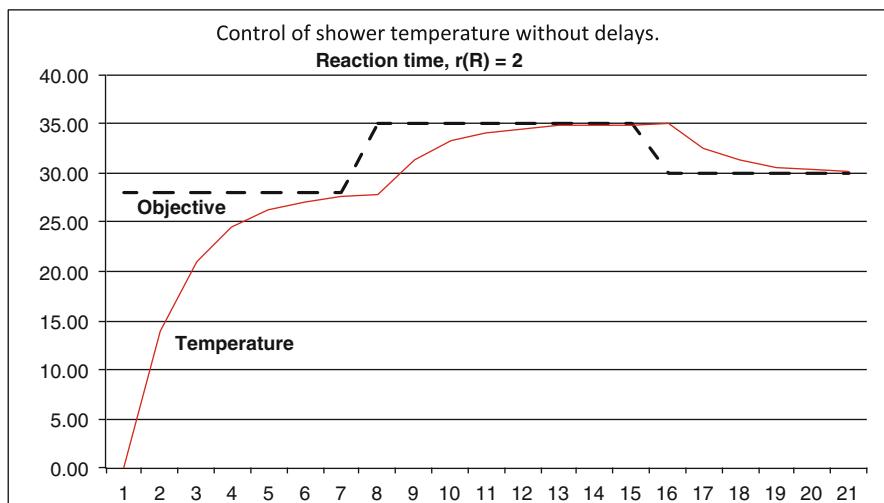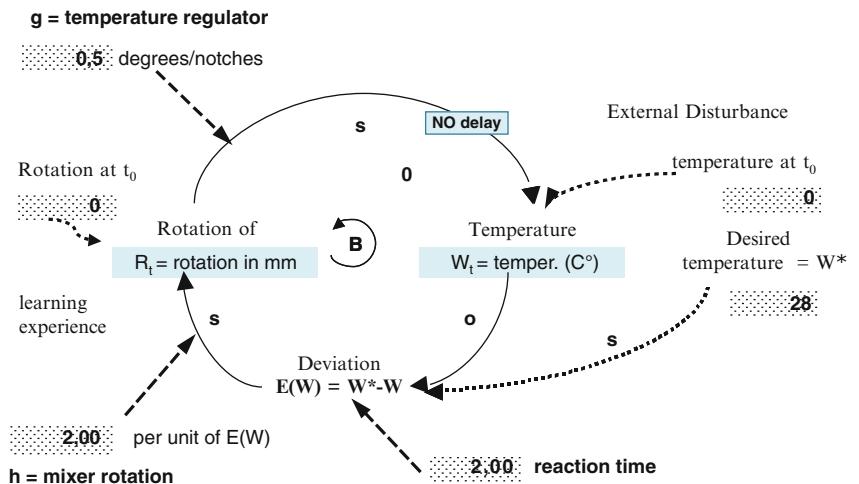
Fig. 2.8 Shower temperature control, without delay

varies the water temperature ($W = Y$) according to the constant rate "$g = 0.5$ degrees/notches," and that the system is perfectly symmetrical, so that "$h = 2$ notches/degrees." Let us assume the objective is a temperature "$W^* = 28°$" and that, as the shower is symmetric, for each degree by which the temperature *differs* from the objective the regulator must rotate the mixer 2 notches ($h = 2$ notch/ degrees) to vary the water temperature by one degree (condition of symmetry); or, what is the same thing, that for each notch of rotation of the regulator the temper- ature rises "$g = 0.5$ degrees/notches."

Figure 2.8 shows that if the system is *without delays* then only a few regulations are needed to reach a stable temperature of $W^* = 28$, even if the manager thinks it

opportune to vary the temperature objective twice (at $t = 7$ and $t = 15$). These assumptions can obviously vary from shower to shower, and the model can easily take account of different values. However, the model in Fig. 2.8 is not realistic. How many times, in fact, have we turned on the mixer and not felt the expected variation in temperature, so that we have to turn the mixer some more, with the devastating effect of ice-cold or boiling water. The water temperature control system normally has at least one *delay*.

Here are some difficulties we may encounter in controlling the shower temperature when there is a delay: we are under the shower and, feeling cold (this is the signal that accompanies the "error temperature"), we turn the mixer (initial impulse) to increase the temperature. Since the water has a rather long route to follow, from the boiler to the shower, some seconds, "Δt," are required for the temperature to adjust to the rotation of the mixer. The cold is unpleasant and we are not willing to put up with it, so we turn the regulator further (second impulse); finally, after an interval "Δt" (which depends on the physical structure of the shower, the length of the tubes that bring the water, the wall temperature, etc.) the temperature turns pleasant, as a response to the initial rotation of the mixer (initial impulse).

However, after another interval the effect of the second impulse kicks in and the water begins to scald us. We rush to reverse the rotation of the regulator but nothing happens, because the delay also operates in reverse; so there is a new impulse, with the result that we might produce too strong a variation with a consequent discharge of ice-cold water, as we have all experienced. If we do not "jump out" of the shower or turn off the water, but instead patiently continue regulating the mixer, we usually end up with a pleasant temperature. This phenomenon occurs most often in hotel rooms when we shower for the first time; thereafter the regulation is easier, because we have learned how the water temperature "responds"—in gradations and with a reaction *delay*—to the rotation of the regulator.

Figure 2.9 shows the dynamics of the values of the same shower in Fig. 2.8, in which I have assumed *a single delay*.

The first graph indicates the dynamics of W_t and R_t under the assumption of a *reaction time* of one period: that is, an instantaneous adjustment of the temperature following the rotation of the mixer. The second graph presents the dynamics for the same variables with a *reaction time* of two periods.

We immediately observe that, while the instantly adjustable shower (first diagram above) does not allow us to achieve the desired objective, but leaves the poor manager–user continuously going from hot to cold, the presence of a *reaction time* of $r(R) = 2$ (second diagram down) makes the water temperature adjustment slower and more gradual; the oscillations are less accentuated and quickly "attenuate" to allow the objective to be reached. This strange behavior shows that delays make it difficult to set up an automatic control that tends to reach and maintain an objective through a mechanical regulator. However, the different dynamics show us why, despite the fact all showers always have some built-in delay, we are always able to take a shower: the presence of a reaction time $r(R) > 1$ produces a gradual adjustment of the temperature that makes slight oscillations "acceptable" while we wait for the desired objective.

	W*	28				g =	0.5
	W at t₀ =	0				h =	2
	R at t₀ =	0				r =	1.00
						tolerance objective	0

t	Rotation of mixer $R = R - (E * h)$	temperature regulator = g	$W = R *g$	D = disturbance	desired temperature =W*	variation in objective	Error $E(W = (W*-W)+D$	mixer regulator = h
0	0.00	0.50	0.00		28.00		28.00	2.00
1	56.00	0.50	0.00		28.00		28.00	2.00
2	112.00	0.50	28.00		28.00		0.00	2.00
3	112.00	0.50	56.00		28.00		-28.00	2.00
4	56.00	0.50	56.00		28.00		-28.00	2.00
5	0.00	0.50	28.00		28.00		0.00	2.00
6	0.00	0.50	0.00		28.00		28.00	2.00
7	56.00	0.50	0.00		28.00		28.00	2.00
8	112.00	0.50	28.00		28.00		0.00	2.00
9	112.00	0.50	56.00		28.00		-28.00	2.00
10	56.00	0.50	56.00		28.00		-28.00	2.00
11	0.00	0.50	28.00		28.00		0.00	2.00
12	0.00	0.50	0.00		28.00		28.00	2.00

Fig. 2.9 Shower temperature control: 1 delay and 2 alternative reaction times

Our simulation model applies naturally to non-automatic systems as well, even though it is obvious that a manual regulation to adjust the shower temperature would have required a different number of attempts based on the manager–operator's experience.

This *conclusion* is general in nature: as the *existence of a delay* causes serious regulation problems, since it makes it difficult to reach and maintain the objective, the manager–operator of the system develops a *learning process* that, by learning from errors and delays, makes the control more effective. A second *conclusion* is also clear: in order to describe and simulate a single-lever control system with delay, not only is it necessary to specify the action "g" and reaction "h" rates, the reaction time "r," and the *initial* values of both X and Y, it is also indispensable to determine the number of delays that intervene between a variation in X and the corresponding variation in Y.

2.9   The Technical Structure of a Single-Lever Control System: The Chain of Control

Despite their apparent simplicity, the control system models illustrated in previous sections have allowed us to make some general observations. Those models need to be further developed, but even the basic ideas presented so far allow us to understand the full power of control systems and the fact that without the concept of *control process* and without control systems as instruments for realizing such a process, we cannot fully understand the dynamics in our world.

One thing is certain: without the action of billions upon billions of *Rings*, of varying hierarchical levels, life itself could not exist, neither single-celled life nor, even more so, the life of organisms, man as an individual, organizations, society, and ecosystems, as I shall demonstrate in Part 2 of this book.

In order to fully grasp the action of control systems we must always remember that the dynamics of the variables shown in the model in Fig. 2.2 are *produced* by interconnected *machines*—that is, by physical, biological, social "apparatuses," or those of other types and configurations—that develop the processes on which the dynamics of the variables depend. Through their interconnections these *machines*, whatever their nature, form the *technical structure* (or *real system*) on whose processes the *logical structure* (or *formal* system) of the control system depends.

By integrating the *logical structure* presented in the standard model of Fig. 2.2, this chapter proposes to define the *general technical structure* of control systems. Figure 2.10 shows a *complete model* which applies to any single-lever control system since it includes the "apparatuses" that constitute the *technical structure* of control systems, which is usually investigated by engineers, biologists, physiologists, economists, businessmen, etc. Knowledge of the *machines* and *processes* that produce the variations in X_t and Y_t is fundamental; unfortunately, in many cases their nature and behavior cannot be observed or understood, in which case they are considered *black boxes* and are not represented in the model; in such cases only the logical model can be represented.

We can generalize: all control systems depend on a *technical structure* made up of three basic process-producing "machines" linked together by an *information network* (not shown in Fig. 2.10) which processes the information (mechanical,
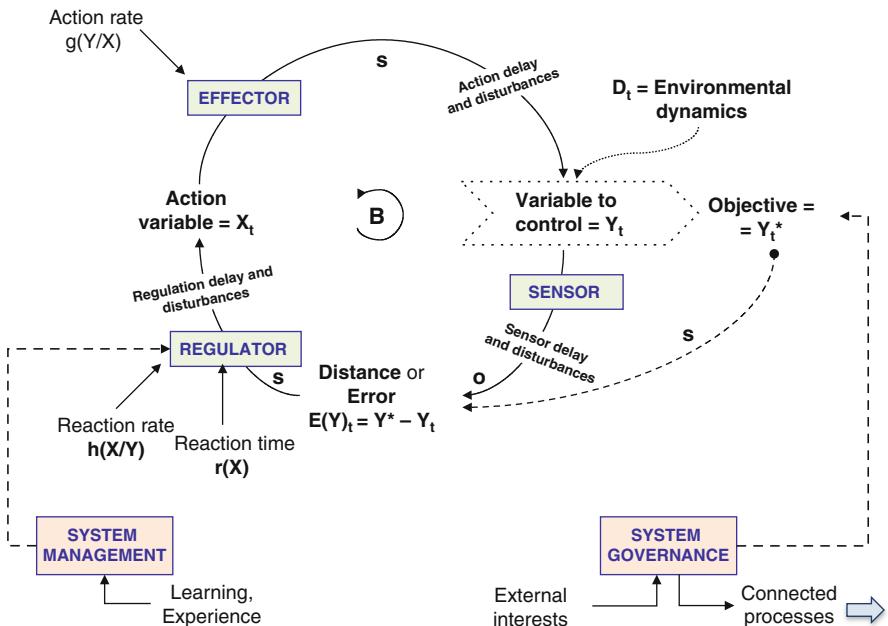
Fig. 2.10 Complete model of a single-lever control system

electrical, electronic, human, etc.) necessary for the functioning of the apparatuses, transmitting this information to them.

(a) The *effector* (or *operator*) which represents the *apparatus* (natural or artificial) which, in fact, enables a variation, ΔX_t, to be transformed into a corresponding variation ΔY_t, taking account of the *action function*, "$g(Y/X)$," and the values of the disturbance variables "D_t"; according to the engineering logic of the control (Sect. 2.16.10), this apparatus is considered as a physical system to be controlled.

(b) The *sensor* (or error *detector*) which represents the *apparatus* (natural or artificial) whose function is to quantify the deviation or error by measuring the value of Y_t in a form useful for comparing it with the objective $Y*$ and determining the error $E(Y)$; without a precise calculation of the error we cannot regulate X_t; from a technical point of view, the sensor can be thought of as composed of a *detector*, the apparatus that measures the values of Y_t, as well as a *comparator* which, through appropriate mechanisms (transducers and actuators), compares at each moment the values of Y_t with those of $Y*$, calculating the deviations E_t.

(c) The *regulator* (*controller* or *compensator*) which represents the *apparatus* capable of "operating" the action variable—that is, calculating the new value of X_t, taking account of $E(Y)$, based on a *reaction function* "$h(X/Y)$"; technically the regulator is the controller apparatus in a strict sense; by using the data provided by the detector it acts on the effector.

(d) The *information network* which is made up of various types of information transmission channels linked in various ways and which transforms the three apparatuses above—together with the *processes* carried out—into a "real" control system. This transmission network is not shown in Fig. 2.10, since it can be considered an integral part of the three "machines" that make up the real *control chain*.

Note that the concepts of network and *information transmission channels* concern not only the mechanical, electrical, electronic or biological links among the apparatuses in the control chain; they must also be taken in a wider sense as a component of the technical (physical) structure of control systems—taking into account the various types of such systems. If the connecting chain (information transmission apparatus) between the rudder wheel (control lever) and the rudder (effector) should break, then it would be impossible to control the ship's direction; similarly, a trauma that damaged the nerves in a part of our body (information transmission apparatus) would impede the action of the muscles (control lever) and inhibit any postural control by our body.

If the shower temperature (Y) is kept at the desired value (Y^*), this is due to the control system formed, for example, by the skin (sensor), the manual maneuvering (regulator) of the mixer tap (control lever), and the time needed for the hot water heater to turn on (effector), on the condition that the nervous system correctly transmits the sensorial impulses. If the stereo volume (Y) does not exceed the levels that allow us to pleasantly listen to music (Y^*), this is because of a control system composed of the ear (detector), the regulation of the sound level (regulator) by rotating the knob (control lever), and the internal apparatus that produces the sound (effector), on the condition that the network of electrical connections allows the apparatuses to function properly. If the quantity of light reaching the retina (Y) is relatively constant throughout the day (Y^*), this owes to the action (control lever) of a muscular apparatus (effector) that modifies the opening of the pupil (effector), with the eventual aid of eyeglasses and artificial light (voluntary control levers adopted by the manager). If our body temperature (Y) is maintained within its physiological limits (Y^*), this is due to a control system made up of the skin (detector) which transmits the sensation of hot/cold (error) to the brain (regulator) so that it can regulate the temperature through the biological apparatuses (effectors) of sweating and chilling (control levers), aided by clothes, heating, cooling, and transferal to environments with favorable climates (effectors).

No part of our world can exist without a control system, in the absence of which the variables take on unacceptable values and a catastrophe occurs; the system that variable is a part of is disturbed and even destroyed. It is all too clear that when arms escalation and increased violence occur, the tragedy of war cannot be avoided without the help of a control system. Nevertheless, no war drags on forever, since sooner or later a political and social control system is created that stops the escalation. Control systems are often difficult to perceive since they are so microscopic (control systems in cells, the eye, neurons), vast (control systems for climate, population dynamics, society), or complex (control systems for neural activity, the immune system, the economic system) as to make their representation and understanding difficult.

There is no doubt that if man has survived for millennia and attained his present level of prosperity, this owes to his success in controlling the waters, fertility, productivity, epidemics, and wars. As yet no effective control systems for births, global temperature, desertification, typhoons, and new epidemics have been conceived, but everything leads us to believe that such systems will be devised and constructed as soon as effective apparatuses that include these phenomena are devised.

First technical note. Figure 2.10 presents a general model of a control system where the control lever, X_t, acts directly on the variable to control, Y_t. This represents a so-called perfect *Ring*, since the three fundamental variables, $(X_t, Y_t$ and $E(Y)_t)$, are directly linked through a single complex "apparatus". The FIRST GENERAL RULE of systems thinking teaches us that if we wish to construct useful systems models we must develop the capacity to "zoom" from the whole to the parts, from systems to components, and vice versa. We can apply this rule to control systems as well; if we want to have an in-depth understanding of how one variable acts on another, we can zoom in and break up the "apparatuses" into smaller components, thereby specifying the links between them.

Second technical note. Normally we assume a control system is formed by an *unvarying* structure when the apparatuses that make up the *chain of control* are characterized by action parameters whose values are unequivocally determined and do not vary as a function of X, Y, or $E(Y)$. I shall define a *varying structure* control system as one whose structure (and thus capacity for control) is varied by the apparatuses as a function of one or more of the systems' fundamental variables. Operationally speaking, in control systems with an *unvarying structure*, $g(Y/X) = g = $ cost is an action *rate*; if, instead, we have an action *function* $g(Y/X) = g(X, Y, E, D)$, then the system would have a *varying* structure and the effector might even produce a nonlinear dynamics in Y_t as X_t varies. The possibility of using the apparatuses in the chain of control to vary the action function usually derives from the system manager's ability to strengthen, when necessary, the power of control, as we shall see in Sect. 5.4 (Fig. 5.2).

2.10 Management and Governance of the Control System

Figure 2.10 indicates two fundamental actors that in various ways normally intervene in any control system: the *manager* and the *governor*.

Above all, I would like to reflect on the *regulation mechanism* and *process* whose function is to decide which new value (or variation) to give the action variable, X_t, taking into account the amount of error, $E(Y)$, in order to enable the system to drive Y_t to reach Y^*. The *regulation process* is carried out by an actor I shall define as the *system manager*. More precisely, the *manager* of the control system (in the broadest sense of the term) is defined as the subject (individual or group, organ or organization) that, by means of a *series of decisions*, operates on the *regulator* to vary X_t in order to modify Y_t, assuming that the functioning of the *apparatus* does not vary over time. The *manager* must almost always directly

intervene on the *effector* and on the *sensor apparatus*, taking into account the values of the functions $g(Y/X)$, $h(X/Y)$, and $r(X)$.

Precisely for this reason the *manager's* activity—the *management process*—is based on appropriate knowledge (together with sufficient experience) of the entire *control chain*; such knowledge must allow the manager to judge the *appropriateness of the objectives* to be achieved, to carry out the *information-gathering* procedures to determine the error and quantify the *delays*, and to grasp the functioning of the *effector*, as well as to gain mastery over the operational mechanisms of the *regulator* and the *information network*.

A *cybernetic system* (or a system with *internal* regulation) is a control system whose *manager* is an integral part of the chain of control. No management intervention in a cybernetic system comes from outside. The system is self-sufficient in producing its own dynamics in order to achieve the objective. If the manager was instead outside the chain of control, separated from the regulator, then the system would be *externally regulated*.

Under a normal shower we are the *managers* of the control system, detecting the temperature with our skin and regulating the mixer lever with our hands, even if we cannot control the boiler that heats the water, the length of the tubes it runs through, or the external temperature. If we include within the chain of control interventions by the user–manager as well, then a normal shower becomes a typical close-looped control system; to an outside observer it would be considered a cybernetic system in all regards. On the other hand, if we use a modern shower, which allows us to set the desired temperature on a thermostat, then we have a cybernetic system, and we are only users but not managers of the system. If the control levers are external to the shower and, when we want to vary the water temperature, we have to ask an outside manager to regulate the mixer, then the control system cannot be defined as cybernetic, even though it maintains the features of a control system with feedback. Similarly, when we drive we are the internal managers that control our cars, since we influence a certain number of action variables (brakes, gas pedal, gear shift) in order to regulate speed (variable under control) so as not to exceed an external limit (constraint). For an outside observer, on the other hand, a car with its driver (for example, the taxi that must take us to our destination) is in all respects a cybernetic system.

In our offices the temperature is automatically regulated by a thermostat whose desired temperature (objective) we determine; this is an automatic cybernetic system. The cook who regulates the gas to maintain the oven temperature stable is an external manager; in order to avoid burnt or undercooked cakes, kitchens have for some time now been equipped with an automatic control consisting of internal thermostats that signal an error in temperature to a regulator which then varies the flow of gas. We can thus reach the following conclusion: a system is "cybernetic" *to the outside observer* when the manager of the system is an integral part of the chain of control, so that the system tries to achieve the objective set by the *governance* without interventions from outside control.

Referring precisely to control systems with feedback, an *initial question* arises: if the manager is part of the control system (internal or external), then who sets the *objective*? I shall define *governor* as the actor (individual or group, organ or

organization) that sets the objective, Y^*, which management must achieve. Thus the governor is not part of the control system but an *external* agent that considers the control system as *instrumental* for achieving the objective Y^*, which is necessary for him/it to achieve his/its own personal interests or produce other *processes down the line*, as shown by the arrow in Fig. 2.10 leading from the box labeled "governance" (lower right of figure). The control system for the order of finish of the competitors in a race is instrumental for the regularity of the race itself; the control system for the production specifications of a machine is instrumental for obtaining the desired quality components to be used in the production process; the control of the stereo volume or shower temperature is necessary in order to allow oneself several minutes of relaxation. The *governance* process can also be complex, particularly in the many cases where the system is viewed as *instrumental* by various actors at the same time, each of whom proposes a different objective in order to attain its own individual interests. Thus the *system objective* is defined through a "political negotiation" among the actors on the basis of the latter's relative power. We know this all too well from having to reach an agreement with our office colleagues about the temperature to set on the thermostat, or when we are driving other passengers and have different opinions on the proper speed, or when we have to discuss with other family members the best place to go for our picnic.

2.11　Design and Realization of the Control System

A *new question* now arises: who in fact constructs the control system to achieve the objective set by the *governance*? To answer this we need to consider two new actors external to the system: the *designer* and the *constructor*, understood in the broadest sense. These are the actors who, after assessing the objectives, identify the most suitable chain of control structure and determine the functions $g(Y/X)$, $h(X/Y)$, and $r(Y)$.

The *manager* limits himself to making a certain number of decisions involving the regulation of an already-formed chain of control (whose functioning he probably is not familiar with) that is developed by the *designer* and built by the *constructor*, both of whom are external subjects, concrete or ideal, who make the system to control available to the manager.

There is no doubt that we are the managers of the car we are driving through city traffic or that our car is the product of the engineers that designed it and the firm that built it. Today the car industry is designing cars with automatic controls which, by means of systems for determining position through General Position Systems (GPS) *devices*, will enable a destination (spatial objective) to be reached in the most efficient way possible through an increasingly integrated control of the vehicle (see Sect. 6.2).

It should be noted that it is not easy to design and construct a control system. The designer and constructor must take into account all the factors that can alter the values of Y_t. Some of these are *external*, being linked to "environmental variety," while others are internal and innate to the "endogenous variety" that characterizes the technical functioning of the apparatuses making up the system's physical structure. For example, it is clear that a control system for an airplane's speed must function whether the

plane is flying through calm air or in the midst of turbulence, near sea level at 20 °C or at 36,000 ft at −30 °C, etc. (environmental variety); or flying on a full or semi-full tank (the weight decreases along with the distance travelled); with efficient engines or less efficient ones, etc. (endogenous variety).

I would like to point out that most systems that operate in our body to keep us alive are automatic control systems. Among the simplest and most evident are those that regulate our heart beat, or arterial pressure, in order to adjust this to the movement of our body so as to maintain a constant supply of blood; those that control body temperature through sweating or chills; those that ward off fatigue by requiring the body to rest; and those that regulate the intensity of movement and generate tiredness. It is equally clear that such systems must function efficiently, both in young and not-so-young people (endogenous variety), in warm or cold environments, at sea level or in the high mountains (environmental variety).

At times physiological control systems are not only completely automatic. We become the external manager of a chain of control; if sweating is not sufficient to lower our body temperature we take off our clothes; if we still feel hot we take a cool shower; and if the heat persists for a long time we take a holiday in the mountains.

2.12   Delays and Disturbances in the Control: Variants of the Control Model

The model in Fig. 2.10 also shows us that in the *logical structure* of control systems delays are inevitable, since they occur in the *technical structure*.

> Delays between actions and consequences are everywhere in human systems … Unrecognized delays can also lead to instability and breakdown, especially when they are long (Senge 2006, p. 88).

> If a government manages to cut back on births today, it will affect school sizes in about 10 years, the labor force in 20 years, the size of the next generation in about 30 years, and the number of retirees in about 60 years. Mathematically, this is very much like trying to control a space probe far out in the solar system, where your commands take hours to reach it, or like trying to control the temperature of your shower when there's a half minute delay between adjusting the tap and the hot water reaching you. If you don't take the delay into account properly you can get scalded. (William Brian Arthur 1973, in Waldrop 1992, p. 27).

There are three types of delay (Coyle 1977; Roberts et al. 1983):

(a) The most frequent is the *action delay*—or output delay—which depends on the effector and slows down the response of Y to an impulse from X; this delay depends on the *effector*.
(b) The second is the *delay in detection*, or information delay; this is a deceptive delay, since it acts on the perception and measurement of the error; if the error is not quickly detected, the regulator can produce a new and dangerous impulse from X.
(c) The third is the *regulation delay*, which occurs when the *regulator* does not promptly respond to the error.

These delays are always lying in ambush. They are inevitable, and only when they are negligible, with a normal standard of tolerance, can the control system be defined as *without delays*. Often we are not aware of their presence because the system admits a certain *tolerance* in its functioning, so that slight asynchronies in the variation of the variables do not noticeably influence its behavior.

Figure 2.10 shows the *disturbances* along with the *delays*, since the former can also create or accentuate the delays. Usually the disturbances act on the variable to control, Y.

> Disturbances: Disturbance is a signal which tends to adversely affect the value of the output of a system. If such a disturbance is generated within the system itself, it is called an internal disturbance. The disturbance generated outside the system acting as an extra input to the system in addition to its normal input, affecting the output adversely is called an external disturbance (Bakshi and Bakshi 2009, pp. 1–3).

More generally, by *disturbance* I mean any phenomenon internal or external to the system which in some way makes the activity of a given apparatus inefficient; in fact, along with delays, disturbances also act on each of the apparatuses that make up the chain of control.

In the examples of the control systems for water temperature and audio volume we can easily recognize the close connection between delays and disturbances. A body which has been in the cold perceives a higher temperature with a longer delay than one that has been in normal conditions (*detector delay*), just as a person who for hours has listened to music at full volume in a discotheque perceives a radio volume that is too high with a delay, thereby angering his or her neighbors. Those who have soap in their eyes and feel the scalding water have a difficult time finding the temperature regulator (*regulation delay*), just as someone listening to music in the dark has difficulty reaching the volume control on the radio.

You have invited friends over and want to amaze them with your homemade pizza. After preparing the pizza with the dough, sauce, and mozzarella cheese, you put it in the electric oven (*effector*) because the instructions say to cook it at 180 °C (objective: Y^*). Even if you are not an expert cook you can regulate the appropriate dial (*control lever X*)—based on your acquired experience—so that it provides a temperature of 180 °C (controlled variable Y). If the oven has an internal thermometer that gives the temperature (*sensor*), the temperature will not cause problems, and you need only check the cooking time (a second variable to control). If the oven does not have an automatic *sensor* for the temperature, or you do not trust the one it has, then you become the *manager* of the system and must periodically visually check (*detecting*) the cooking progress of the pizza. If you see that the pizza is cooking at too low or too high a heat (error between Y^* and Y), you intervene to regulate the dial (*control lever*) several notches to vary the oven temperature in the desired direction. The three delays are waiting in ambush: the oven can adjust the temperature more slowly than planned (*action delay*); distracted by the arrival of your guests, you do not check how the pizza is cooking at frequent enough intervals (*detection delay*); and even though you detect an error you might wait several minutes before regulating the gas knob (*regulation delay*). Making a pizza is not

only a question of the right recipe (productive combination) but also of the friends waiting to have it for dinner (users or clients); above all it represents a *managerial problem regarding the control of a process*.

Delays in detecting the error or in the action of the control levers are often due to physical limits, or limits related to obtaining and transmitting the information. Our car thermometer always shows with a delay the outside temperature that indicates the presence of ice, and we could decide to slow down after having gone some kilometers unaware of the icy roads. Moreover, if we are not careful we might not notice the temperature signal and employ the levers to reduce speed with a delay. For this reason, many cars even adopt an acoustic signal to alert the driver when the temperature is reached at which ice forms. This reduces the risk, as long as we are not driving with the stereo going full blast, which would interfere with our hearing the signal. To control the movement of the Mars probe, it is necessary to detect its position and send signals with information about the new objectives and the tasks to carry out. This requires several minutes between the reception and transmission of the signals, and it is possible that a probe approaching a gulley could receive a message to halt after it has already fallen in.

First technical note. Although delays are linked to the time it takes for the processes to be carried out, these should nevertheless be identified or confused with the *normal* functioning times (even when lengthy) of the apparatuses carrying out those processes. The delay must be considered as such by management when it perceives an apparatus response time *greater than what would be expected under normal operating conditions*. We all know that by turning the mixer faucet we do not produce an instantaneous variation in temperature; if our experience has memorized an action time of, for example, 4 s, then we will start to perceive a real delay after 6 or 8 s, which leads us to turn the dial further. A delay beyond the normal functioning times can result from our error in maneuvering the levers, from a functioning disturbance in the apparatuses, or even from external events. Only delays should be considered as anomalous, and there is no remedy for them except to trust in our experience and attentively and regularly observe the environment.

Second technical note. Since delays are always held to be a destabilizing factor, they are the main focus of the analysis and design of control systems. However, we must not underestimate the effect on control systems from the opposite situation: a *premature occurrence* that is manifested when apparatuses respond in a shorter time than management expects under normal operating conditions. Whether or not Y_t varies earlier than our experience indicates is normal or with a delay, our reaction is the same: we try as soon as possible to reset the values on the control lever, X_t. We can deduce that a premature occurrence produces control problems which in a certain sense are the same as those resulting from a delay. In fact, a premature occurrence forces the manager to reduce the action and reaction times with respect to the normal (longer) ones that regulate the action of the technical apparatuses. For the manager it is as if the system responded "with a delay" to the premature action of the levers.

2.13   Strengthening and Precision of Control Systems

Let us return to the *general model* in Fig. 2.10 in order to deal with a technical topic that will allow us to reflect on the historical evolution and the evolutionary tendencies of control systems, while also making some brief considerations that will be further explored in subsequent chapters.

It seems that the most relevant progress in control systems, of whatever type, has come about through the improvement in the power of the *effectors* that control the variable Y through the control lever X (these considerations also hold for the multi-lever systems we shall examine in Chap. 4). There is clearly a greater capacity of control of the route of a boat that has powerful turbine motors equipped with stabilizers controlled by electric motors than one powered by the wind and steered by a manually operated rudder.

But while modern devices that produce the effect of X on Y have considerable power and incomparable precision, the greatest progress in control systems must be attributed to improvements in *detection* and *regulation* apparatuses.

Who tunes his radio or television today by means of manually operated dials and knobs, as was the case at the dawn of radio transmission technology? Nobody thinks twice about this, since now there is automatically and electronically controlled tuning, which is much more precise and faster. From regulating routes using an oar we have evolved to the bar rudder, the circular rudder, and then to the modern rudder with servo-control. Today at home we no longer measure our child's temperature by placing the palm of our hand over his forehead but by using modern laser ray devices that do not even have to be in contact with the skin.

The lack of some devices in the chain of control prohibits us from controlling a number of variables that we could instead influence. While city authorities try to keep under control the amount of fine dust in the air through measurements at a detection facility and by banning traffic, we have yet to find appropriate effectors to control global warming, even though this phenomenon can be measured with precise instruments. It is not science fiction to imagine a future with powerful suction devices that can suck in icy air from the higher strata of the atmosphere to bring to the warmer zones to lower the temperature, or powerful purifiers that could eliminate air pollution in cities. And how will it be possible to control viral epidemics if we lack detection systems to measure the epidemic situation and effector systems to act against the spread of the virus? How can companies activate the market control levers or those for the demand for its products if the marketing departments do not have the proper instruments for detecting the main market variables on which to intervene? These few examples should be enough to give an idea of the progress of control systems, regarding both the precision of error measurements and the development of true control devices.

In the area of mechanical and biological controls we seem to have reached a state of absolute precision: zero tolerance with respect to the functioning, pollution, and maintenance objectives. Evolution in these areas has involved a multitude of attempts at making more efficient apparatuses through which organisms try to control their main processes.

Even a quick survey reveals the evolution in the limbs of terrestrial animals; that is, in the effectors that produce speed and direction. We can also observe the many forms of propulsion in marine animals—from tail fins to forceful water expulsion—and of direction in the water—from rear to lateral to front fins. Birds and many insects are equipped with three-dimensional wings to move in space, and all reptiles have developed a totally or partially crawling movement.

Even the *effector organs* to grasp and capture objects have evolved in quite different ways. The hands, along with the bones and muscular apparatuses of the neck and arm, are the effectors *par excellence* in man and in many primates, while the trunk enables the elephant to grasp things with extreme precision. In order to capture insects by eliminating the distance from their prey, frogs, toads, and chameleons can quickly flick out their tongues. The woodpecker also sticks out its tongue to capture insects inside tree trunks, and we find similar effector organs in mammals that feed on ants or termites, such as the anteater. Many classes of mollusks and polyps have developed very elaborate tentacles they can use to grab prey, many insects, and crustaceans (for example, ants, beetles, lobsters, and shrimp) are equipped with more or less efficient claws to capture their prey.

What variety of *sensors* has evolution produced in nature? In addition to the five human senses of which we are all familiar, evolution has produced many other types of sensors. A very efficient one is the biosonar of bats and dolphins and porpoises, which enables them to perceive the dynamics of many environmental variables and determine the gap between their position and an object. Even Lorenzini's ampullae, which function as electroreceptors of the prey's magnetic field, and the lateral line system, employing neuromasts, are precise receptors that allow sharks and torpedoes to locate the position of external objects and determine the distance to these objects. I will also cite the fantastic world of insects: ants have senses for pheromones, through which they can recognize others of their species and determine the presence of a pheromone trail, which they can follow by calculating the distance between this trail and their position. Bees have compound eyes that enable them to locate the hive at a distance of many kilometers; they are also equipped with tactile "sensiles," olfactory systems, thermoreceptors, and hygroreceptors capable of identifying many chemical and physical variables. Many insect species are also equipped with mechanical receptors that permit them to get information about the position and state of various parts of their body and many winged insects have receptors on their heads to measure wind speed.

Even decision-making and control apparatuses have been refined through evolution. Evolution has progressed from strictly reactive automatic regulators in many living beings to the nervous system (which also acts as an information transmission system) and then to the brain, the most evolved regulatory apparatus, and finally the human brain, which has enabled man to gain knowledge and develop language, science, and technology. Through the cognitive power of his brain man has substantially expanded his apparatuses, inventing machinery to produce things more efficiently, to detect phenomena in an extremely precise way, and to support his control decisions. These aspects will be developed further in subsequent chapters in Part 2.

2.14   Connections and Interferences Among Single-Lever Control Systems

In this section I would like to introduce the important new concept of the *connection* and *interference* between single-lever control systems.

Two (or more) control systems can be *connected* in the sense that the dynamics of a system's controlled variable depends not only on the values of its control lever but also on the values of the controlled variable of another system. Two (or more) control systems *interfere* if the connection is *reciprocal*. Thus interference occurs when two or more control systems are interrelated in some way so that the variation in the controlled variable of one system also produces changes in the controlled variable of the other and vice versa.

When the systems interfere with each other, the simple dynamics of each system considered on its own can become complex.

Let us assume we have two single-lever control systems that obey the following equations (see (2.6) and (2.7) in Sect. 2.3):

$$B_n = A_n \, g_1 + C_n \tag{2.10}$$

$$A_{n+1} = A_n + (B^* - B_n) \, h_1; \tag{2.11}$$

and

$$Q_n = P_n g_2 + R_n \tag{2.12}$$

$$P_{n+1} = P_n + (Q^* - Q_n) \, h_2; \tag{2.13}$$

In order to connect two showers and create *interference*, we can consider the first shower as a disturbance to the other and vice versa. We can thus transform (2.10) and (2.12) by assuming that the values of the disturbance variable, C_n, of the first system are a function of the value of the controlled variable, Q_n, of the second system (or, in many cases, of the $E(Y)$), and that this interference is based on a parameter "i_Q"; at the same time we assume that the values of R_n of the second system are a function of the value B_n (or of the error of the first system) according to a parameter "i_B."

Taking appropriate account of the time displacement, (2.10) and (2.12) can be rewritten as shown below, by directly including as disturbances the temperatures of the connected shower, while (2.11) and (2.13) would not change.

$$B_n = A_n \, g_1 + i_Q \, Q_{n-1} \tag{2.10/1}$$

$$Q_n = P_n \, g_2 + i_B \, B_{n-1} \tag{2.12/1}$$

Figure 2.11 shows the dynamics of the temperatures and regulations of two showers in the same apartment which interfere with each other (since they are
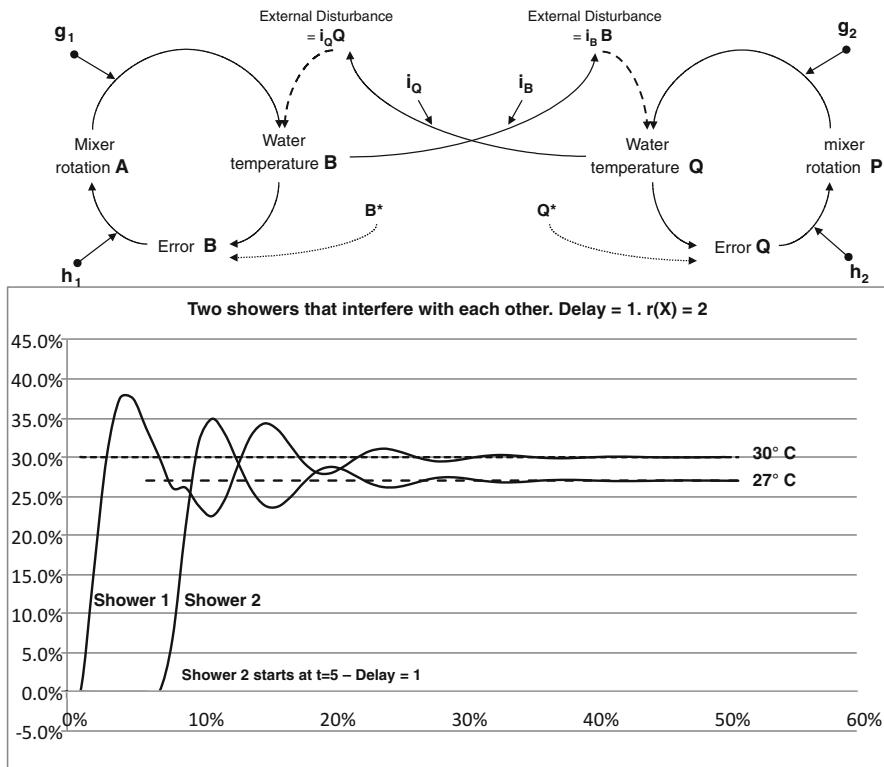
Fig. 2.11 Dynamics of the temperatures of two showers that interfere with each other through Y_t, with $i_Q = -30\ \%$ and $i_B = -20\ \%$ and with delay $= 1$ and reaction time $r(X) = 2$

connected to the same hydraulic system), assuming the same *reaction time* for both showers, $r(X) = 2$.

The first shower is opened at the initial instant; the second begins to function after six periods. To emphasize the interferences I have chosen high values for the coefficients "$i_Q = 30\ \%$" and "$i_B = 20\ \%$."

Equations (2.10/1) and (2.12/1) clearly illustrate the interference process; in fact, we can see that the opening of the second shower produces a lowering of the temperature of the first, which forces the manager of the first shower, in the succeeding instant, to readjust his temperature; however, this causes a variation in the temperature of the second, which is readjusted, only to again interfere with the temperature of the first shower. The dynamics are complicated due to the presence of a *delay* which, in the initial phases, produces an increase in the temperatures of the two showers above their respective objectives. Nevertheless, a reaction time of $r(X) = 2$ protects the control system from breaking down by closing the mixers. A reaction time of $r(X) > 2$ would make the delay even more tolerable.

When the parameters "i_Q" and "i_B" are close to "0" the interferences are modest and the reciprocal variations of the temperatures are not too high, so that the
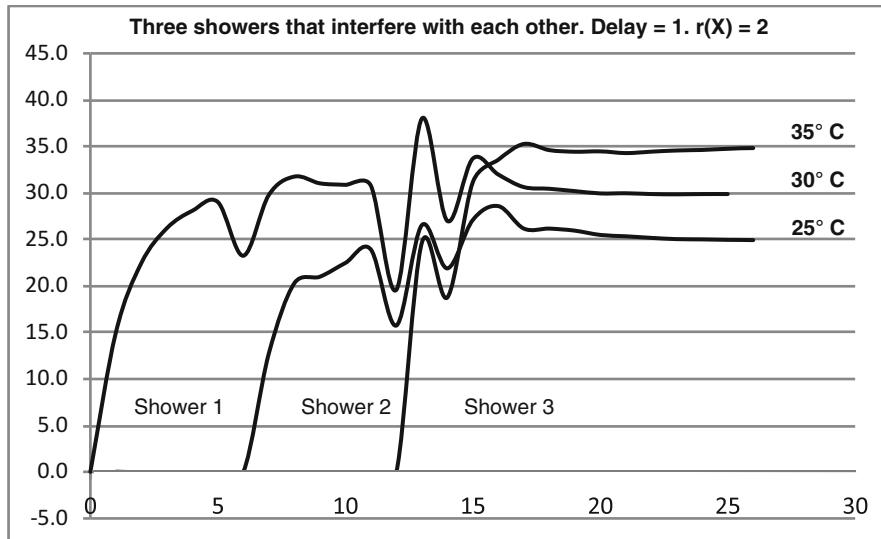
Fig. 2.12 Dynamics of the temperatures of three showers that interfere with each other through Y_t. Delay $= 2$ and reaction time $r(Y) = 2$

managers of the two showers can quickly succeed, with only a few adjustments, to restore the desired temperature. It is even easier to achieve the objectives without any large oscillations by assuming a reaction time of 2 (or even higher).

Figure 2.12 extends the example in Fig. 2.11 by presenting the dynamics of three interfering showers.

In order to work through the example and make the graph understandable, the interference percentages have been set as follows: Shower 1 ($i_2 = -25$ %, $i_3 = -25$ %); Shower 2 ($i_1 = -25$ %, $i_3 = -25$ %); Shower 3 ($i_1 = -25$ %, $i_2 = -80$ %); the symbol "i_k," with $k = 1$, 2, or 3, indicates the percentage at which each shower would interfere with showers to which it is connected. The temperature dynamics of the three showers becomes even more complex by the existence of interference together with the presence of a delay. In any event, in this example all the control systems enable the temperature to reach the objective set by the governance.

Equally complex dynamics would occur if we assume interference quantified not on the basis of the temperature but on the error. In this case (2.10/1) and (2.12/1) would become:

$$B_n = A_n \, g_1 + i_Q(Q^* - Q_{n-1}) \tag{2.10/2}$$

$$Q_n = P_n \, g_2 + i_B(B^* - B_{n-1}) \tag{2.12/2}$$

Figure 2.13 indicates the temperature dynamics of the two showers determined from the last two equations.

Two showers that interfere with each other through E(Y). Delay = 1. r(X) = 2
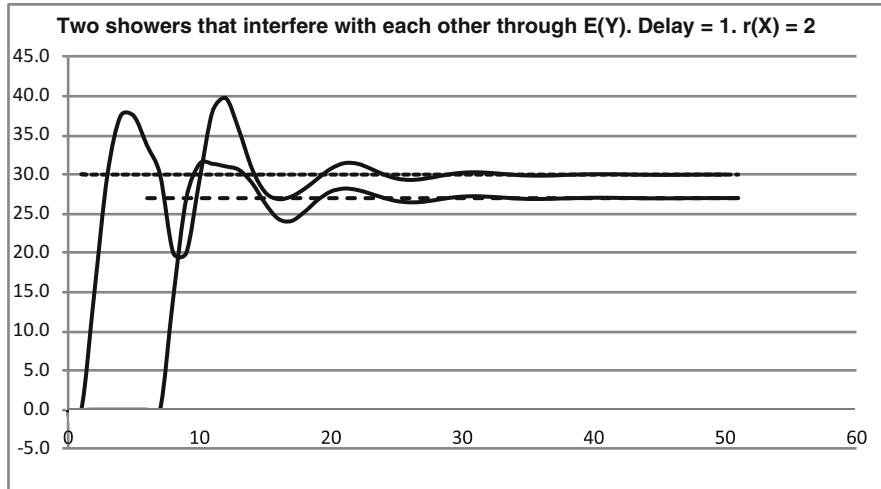
Fig. 2.13 Dynamics of the temperatures of two showers that interfere with each other through E $(Y)_t$. Delay $= 2$ and reaction time $r(Y) = 2$ (compare with Fig. 2.11)

2.15   Areas of Application of the General Model

The general model of a control system in Fig. 2.2 can be found, or applied, in all situations where we wish to modify a variable in order to achieve an objective. It is useful to briefly mention some of these situations, which will be dealt with in more detail in subsequent chapters.

The system we apply in *problem solving* is certainly a control system, as is shown in Fig. 2.14. Solving a problem means achieving a state Y^* considered optimal, starting from a non-optimal state, Y. The problem is perceived as an error between the two states, $E(Y) = Y^* - Y = $ PROBLEM. The problem solver (that is, the manager of the control system) must eliminate the error through decisions involving the activation of the control levers. The model can be applied to both artificial and natural systems.

For example, thirst represents the error $E(Y)$ between an optimal salt/water density (D^*) and the actual density (D); this is the problem we all experience when subject to scorching heat. To get rid of the thirst we must restore the correct saline density by ingesting water (drinking). In all respects we have a control system here that detects the problem (error) and solves it (activation of the lever). This is a natural system for most higher-order animals, who are the managers that decide which levers to activate and to what degree (Sect. 4.2, Fig. 4.6).

Equally clear, in my view, is the representation of the decision-making process as a control system, as shown in the model in Fig. 2.15.

Decision-making represents the process that leads to the choice of some alternative in order to achieve some objective, understood in the broadest sense.
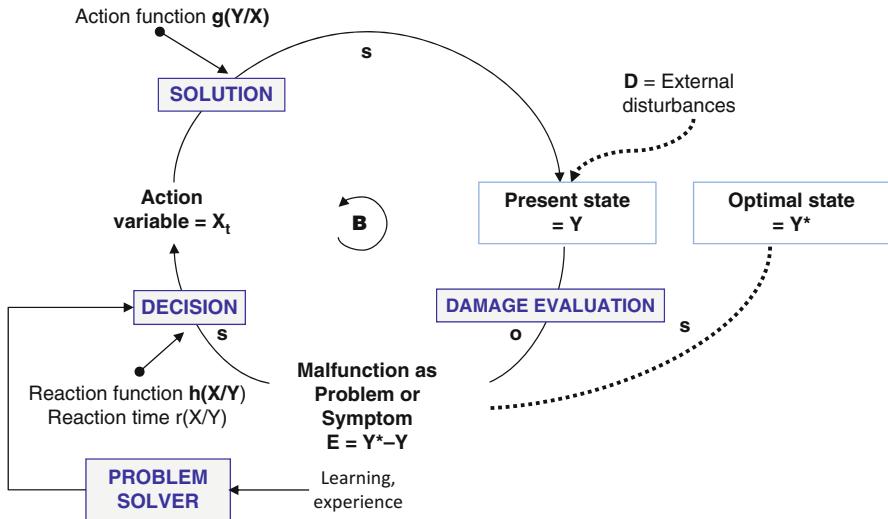
Fig. 2.14 Problem solving as a control system
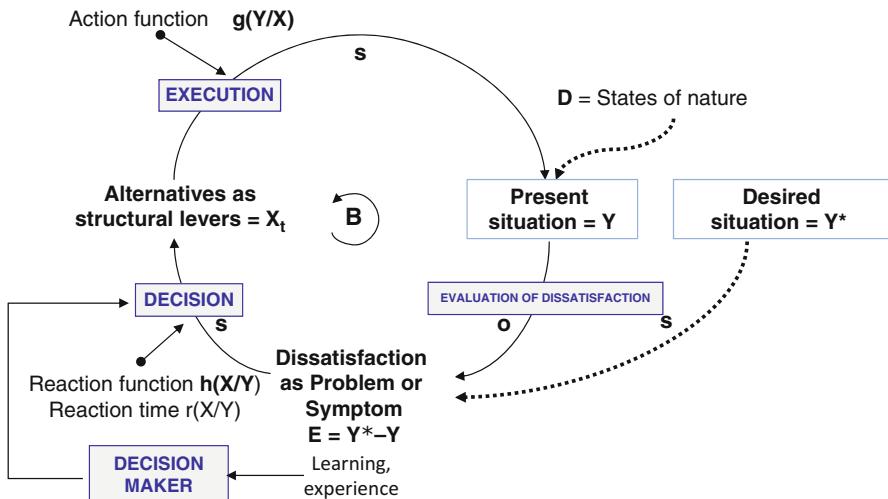
Fig. 2.15 Decision making as a control system

The need to arrive at a decision comes from the perception of a state of dissatisfaction regarding a given situation that contrasts with a situation we hold to be better, and which thus represents the objective to achieve. This dissatisfaction is a symptom of the error between "where we would like to be" and "where we are," and it represents the problem; that is, the error $E(Y)$ to eliminate. From experience the
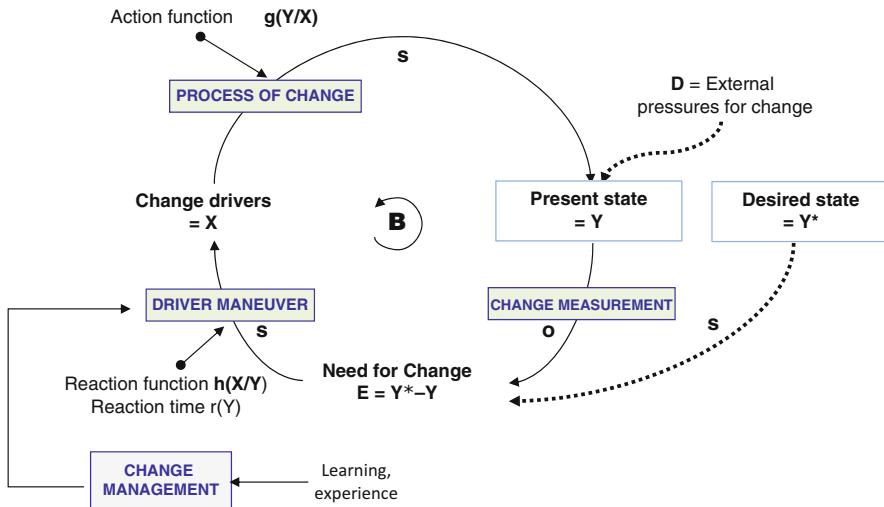
Fig. 2.16 Change management process as a control system

decision-maker identifies, evaluates, and activates the decisional levers, that is, the various action alternatives, to modify the present situation in an attempt to attain the desired one, taking into consideration external disturbances, or states of nature (Sect. 5.8). An efficient information system for evaluating the attainable results and the probabilities of success regarding the various alternatives is activated by the decision-maker while regulating the control levers, X, resulting in the decision-maker choosing the best alternative.

Change is a natural part of the life of every individual, society, or organization. Change requires a modification in one's *present* state, which must be transformed into the *desired* one to be attained; once more the logic of control systems is appropriate, as shown in the model in Fig. 2.16. The management of the control system that produces change can also be referred to as *change management*. Through some *change management* apparatus the need for change is perceived. Change management searches for the *change drivers*; that is, the levers to activate in order to trigger the process of change that should transform the present state into the desired one (Sect. 7.12).

In many cases this change is obligatory, being the consequence of the action of biological programs (physical and mental development) or social ones (learning and education) that define how the change is to come about and at what speed. In many other cases the change is voluntary and motivated by the need to avoid disadvantages or to gain better living conditions. In addition to obligatory change, individuals, organizations, and societies apply decision-making processes whenever they need to produce or face processes of change, whether these processes derive from the need to face environmental situations or are directly activated to produce an improvement in one's situation. These processes of change must be managed so as to maximize
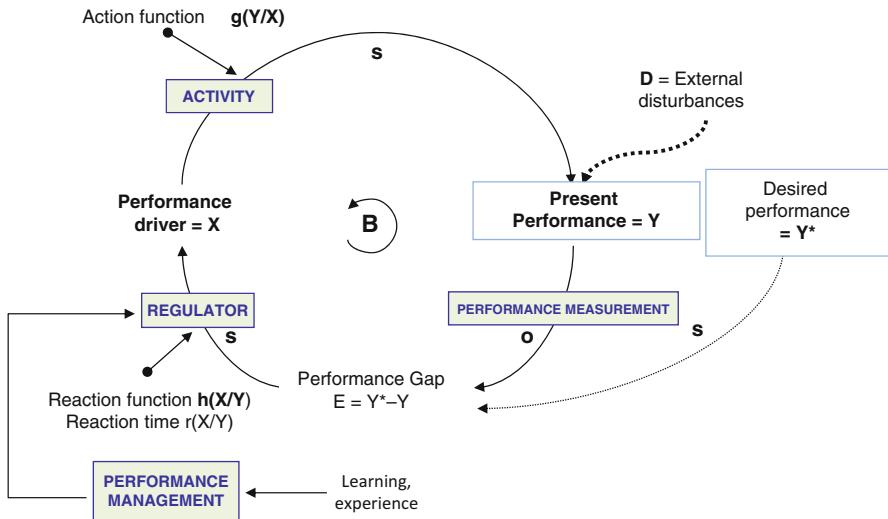
Fig. 2.17   Performance control systems

benefits and minimize costs, and they can be viewed as a generalization of problem solving, if the change comes from *external pressures*, and therefore must be endured, or of decision-making, if the change is actively sought after.

The general model in Fig. 2.10 can also be interpreted as a system for controlling that the performance of any activity is satisfactory, as shown in Fig. 2.17.

When the *performance measurement* reveals a performance gap, $E(Y)$, between the actual performance, (Y), and the desired one, (Y^*), the performance manager intervenes and assigns the most appropriate values to the *performance driver*, X, or the performance control lever, so as to enable the activity to produce a new current performance that will reduce or eliminate the performance gap. The model in Fig. 2.17 has two very general variants depending on the meaning we attribute to the performance concept adopted.

In the first variant, the performance driver X represents the *levels* or *inputs* of an activity or process, while Y indicates the obtainable *results* that must move toward the result objective Y^* by means of the regulation of X. In the second variant the variable X represents the specific techniques of an activity, or process on which performance depends; however, this variant differs from the first in that the variable to control, Y, is not a result but a *performance*, or quality *standard* defined from time to time for the activity or process. For both variants the *performance management* must achieve a desired performance *level*, Y^*, which, in the first case, represents a *performance objective*, while in the second a *performance standard* (Sects. 8.1 and 8.4).

2.16 Complementary Material

2.16.1 Simulation Tools

In Sect. 2.8 the simulation model for the shower temperature using a mixer (Fig. 2.8) was developed using Excel, since a spreadsheet can rapidly and efficiently calculate the necessary iterations to produce the system dynamics. As the reader can observe, even those without a sophisticated knowledge of the topic can develop an automatic control system by simply respecting the rules indicated in the iterative equations model (2.0)–(2.9) of Sect. 2.3, with appropriate adjustments to take into account any operating constraints for the system.

To simulate more complex control systems other software can be used, such as *Mathematics*—the computational program developed by *Wolfram Research*, which has a section specifically dedicated to control systems—or the simulation software based on relations that link *stock*, or *level* variables with *flow*, or *rate* variables, which change the makeup of the stock.

Figure 2.18 illustrates the SFD that corresponds to the example in Fig. 2.8 for the shower temperature control, without delay, with a mixer functioning as a continuous control lever.

From a mathematical point of view, *stock* is a variable that derives from the accumulation or integration of *flows* (with the appropriate sign), while *flows* are variables that add (inflows = input) or subtract (outflows = output) to the amount of the *stock*. The intensity of the *flows* depends on the variation rates (converters) indicated (in various ways) in the corresponding labels. The simulation model that results is called a Stock and Flow Diagram (SFD) or a Level and Flow Structure (LFS). It is obviously necessary to define the temporal scale that determines the dynamics of *flows* and *stocks* (Coyle 1977; Roberts 1978; Sterman 2000, 2001; Mella 2012). Among the many types of simulation software, all of which are relatively simple, we can mention the following:

– POWERSIM (http://www.powersim.com).
– MyStrategy (http://www.strategydynamics.com/mystrategy/default.asp).
– ithink and stella (http://www.iseesystems.com/index.aspx).
– Vensim (http://www.vensim.com).
– Excel Software (http://www.excelsoftware.com/).

I would not be surprised if the reader finds the model in Fig. 2.18 quite different and less significant than the one in Fig. 2.7, which is then simulated in Fig. 2.8, if only because the SFD in Fig. 2.18 does not show the direction of variation and the *loops*, which make the model easy to understand.

In fact, this is the main "limit" of SFD simulation models; but this is balanced by the immeasurable advantage of being able to quickly develop models with hundreds of variables, which is almost impossible with Excel. Nevertheless, if we do not allow ourselves to be confused by the direction of the arrow, we can easily recognize that the SFD in Fig. 2.18 corresponds completely to the CLD in
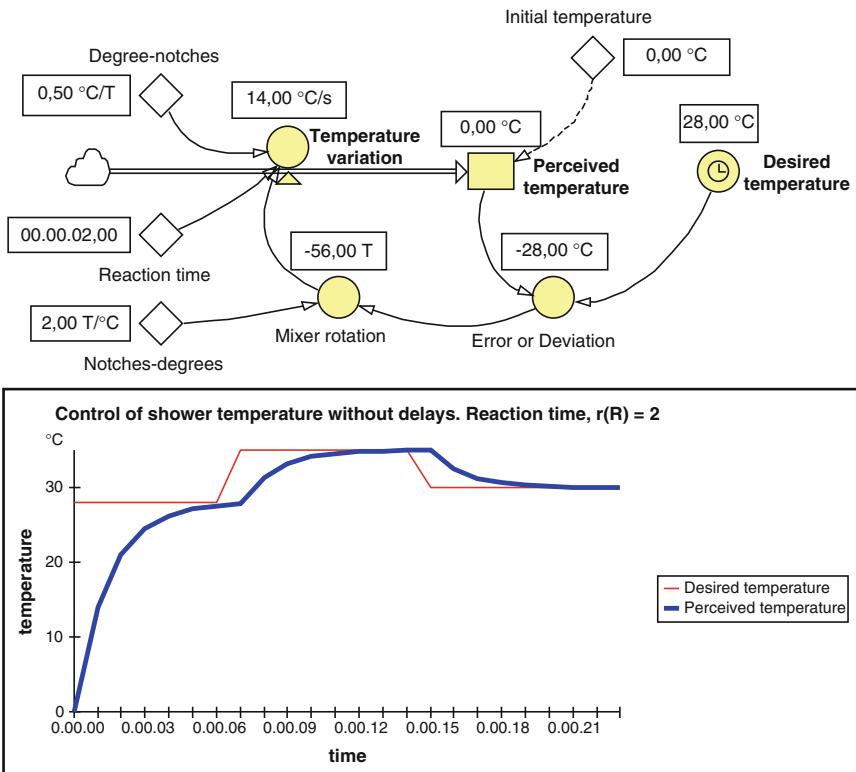
Fig. 2.18 Shower temperature control, without delay, simulated by Powersim (see Fig. 2.8)

Fig. 2.7; in Fig. 2.18 the variable Y_t, which indicates the "perceived temperature," represents a *stock* that is accumulated based on the values of the "temperature variation" variable, which depends on the "mixer rotation" and is equivalent to X_t. All the other elements in Fig. 2.18, in particular the graph showing the dynamics of the temperatures, can be easily recognized in Fig. 2.8.

2.16.2 Control of an Elevator

The simple operation of taking an elevator is made possible by the simple control system in Fig. 2.19.

Figure 2.20 numerically illustrates the dynamics of the elevator, represented by discrete time intervals. When we get on an elevator we want to go to a certain floor (objective), going up or down according to our initial position. To achieve our objective we simply press the button that indicates the floor we want to go to and think no more of it. The pressure from the button indicating the floor is interpreted
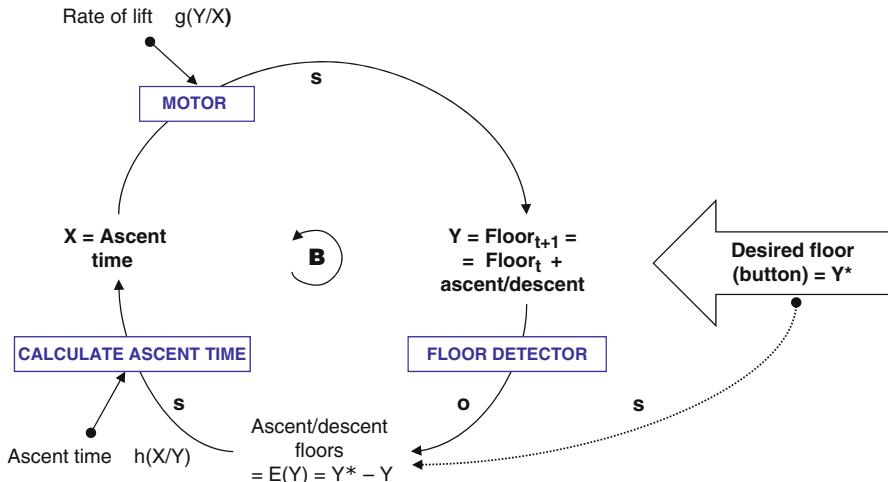
Fig. 2.19   The control system for an elevator

by the machine (effector) as the objective to achieve, determining the displacement relative to the initial floor and turning on the elevator motor for the time necessary to eliminate the gap. The elevator-machine begins to move past the floors that separate us from the one we wish to go to; when the objective is reached, it stops.

2.16.3   Searching for a Street, a Page, and a Word

The functioning logic behind the above-mentioned simple control system for an elevator can be applied to other systems that regulate quite different processes. Above all, the system that allows us to identify a certain street address starting from either end of the street. If, for example, we are at number 5 and want to go to number 11, we only need to behave like the elevator; obviously our trajectory will be horizontal rather than vertical, but the process of eliminating the gap is entirely analogous. We all use the same control system when we look for the page number in a book or newspaper. If we are on page 5 and want to get to page 125, the elevator model is still perfectly applicable.

The model can also be extended to looking up a word in a dictionary. Here we are not controlling the gap between two numbers but that between two words in the dictionary, as shown in Fig. 2.21. As humans we are too used to thinking in terms of letters, that is, symbols; but we should not forget that for computers, letters are merely codified representations of numbers. We can generalize these observations by remembering that as humans we can recognize letters and search for a word in a dictionary because we have learned a catalogue of letters that does not change over time; that is, an alphabet. The recognition of a letter is carried out through a control system that identifies the letters from a catalogue and scrolls from letter to letter,

| 1 Floor | | 1 | $Y_0 =$ | 15 | | |
| | | | $Y^* =$ | 11 | | |

t	X = motor operating time	Y = floor to reach	Y* = Objective	Delta objective	E = disturbance	Ascent & Descent = Y* - Y
0	0.0	15.0	11.0	0	0	-4.0
1	-1.0	14.0	11.0	0	0	-3.0
2	-1.0	13.0	11.0	0	0	-2.0
3	-1.0	12.0	11.0	0	0	-1.0
4	-1.0	11.0	11.0	0	0	0.0
5	0.0	11.0	11.0	0	0	0.0
6	0.0	11.0	11.0	0	0	0.0
7	0.0	11.0	5.0	-6	0	-6.0
8	-1.0	10.0	5.0	0	0	-5.0
9	-1.0	9.0	5.0	0	0	-4.0
10	-1.0	8.0	5.0	0	0	-3.0
11	-1.0	7.0	5.0	0	0	-2.0
12	-1.0	6.0	5.0	0	0	-1.0
13	-1.0	5.0	5.0	0	0	0.0
14	0.0	5.0	5.0	0	0	0.0
15	0.0	5.0	0.0	-5	0	-5.0
16	-1.0	4.0	0.0	0	0	-4.0
17						-3.0
18						-2.0
19						-1.0
20						0.0
[omissis]						

Fig. 2.20 Dynamics of a control system for an elevator

just like an elevator scrolls through its "catalogue of floors." The recognizing systems will be taken up in Sects. 3.1 and 3.9.5.

2.16.4 The Trajectories of a Car and a Boat

The system for regulating the audio level (Sect. 2.5 and Fig. 2.3) can similarly be applied to a vast array of seemingly diverse phenomena. In particular, it is entirely analogous to the control system for the route of any type of car and vehicle that uses a steering wheel to change direction, as shown in the model in Fig. 2.22. The system for regulating a *shower* can similarly be applied to the control system for the route
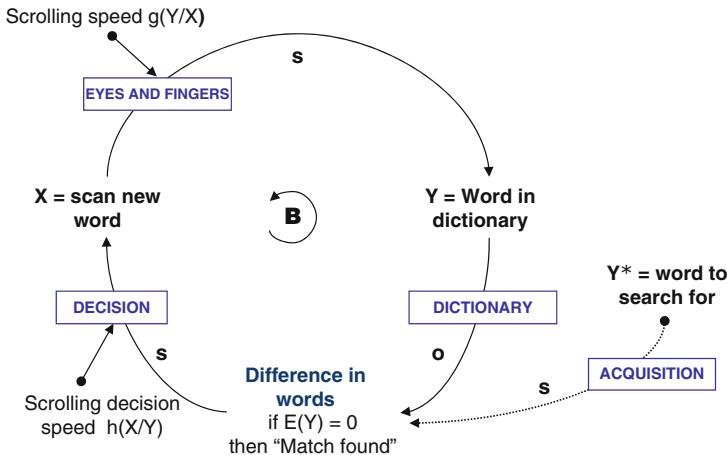
Fig. 2.21   Control system for recognizing a word in a dictionary
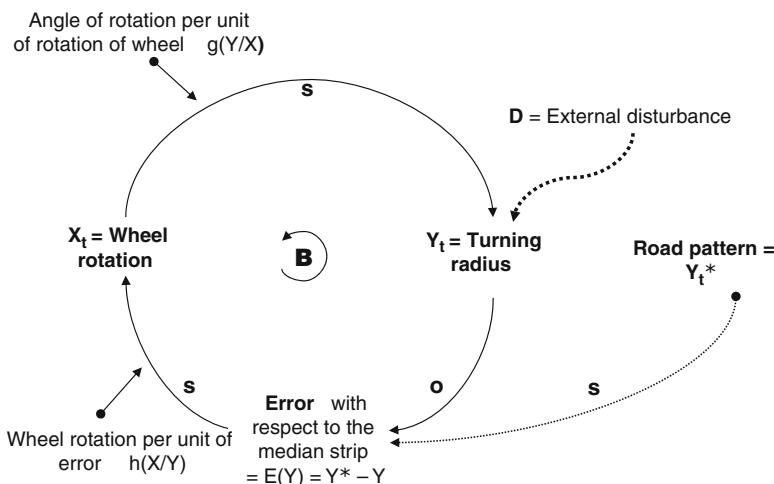
Fig. 2.22   Control system for the trajectory of a car

of any type of boat in a sea or lake (river navigation must also take into account the single-directional current) by means of a rudder wheel. The trajectory of a boat differs from that of a car for two reasons. Above all, the dynamics of a boat is subject to *inertia*; secondly, the direction of a boat is to a great extent autonomous while that of a car is conditioned by the types of curves on the road. In fact, in order to steer a boat toward point Y^*, defined as a geographical location, we turn the rudder a certain number of degrees; the rudder transmits the signal—with appropriate amplifications—to the real rudder, at the stern of the boat, which imparts the appropriate rotation. Since movement in a fluid must overcome inertia, unless we
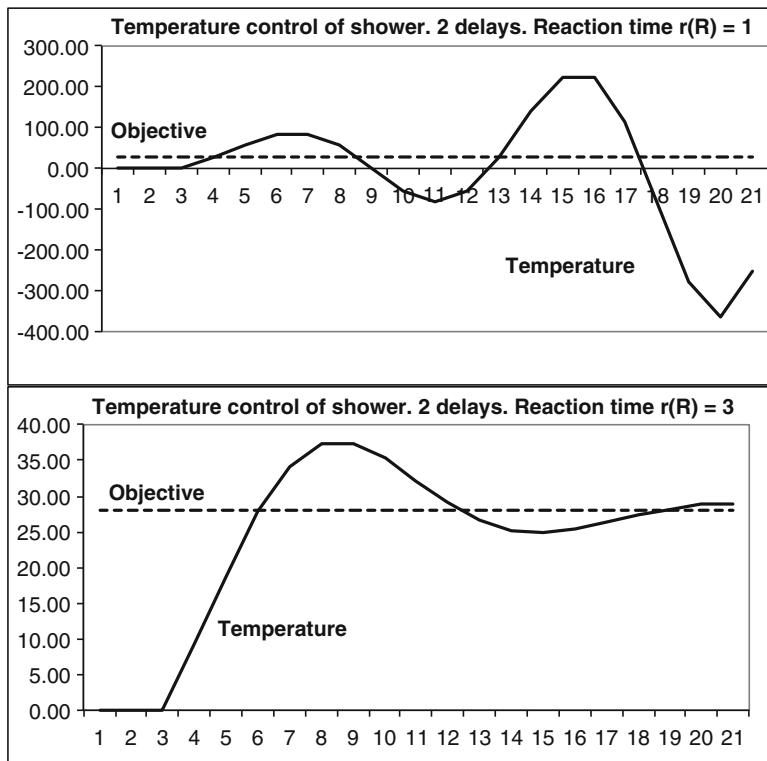
Fig. 2.23 Regulation of a shower with mixer and 2 delays

correctly calculate the rotation of the rudder wheel we run the risk of not achieving the objective.

Modern automatic control systems already take into account *inertia* when determining the correct trajectory; however, to simplify things we can imagine that the control of a boat under the assumption of a manual control is similar to that of the shower temperature when there are *delays*. The helmsman sets the direction but the boat responds with a *delay* due to *inertia*, so that most probably there is a gradually attenuating oscillating movement—entirely like that in the graph at the bottom of Fig. 2.9, with a reaction time of 2 periods—which enables the desired direction to be maintained. The helmsman of a sailboat or a boat without automatic controls must be quite experienced; however, the manual control often leads to failures, as the tragedy of the Titanic reminds us.

2.16.5   Shower with Two Delays

Figure 2.23 presents the same shower as in Fig. 2.9, but assuming *two delays*. Note that in the graph at the top (reaction time of one period) the temperature dynamics inevitably leads the user to "jump out" of the shower almost immediately; the
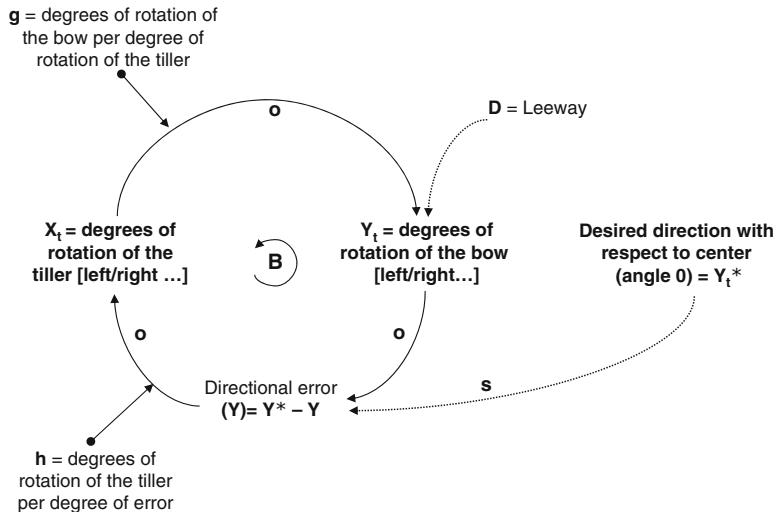
Fig. 2.24   Inverse control system [**o-o-o**]. Control of a boat's direction

reaction time of three periods means the 28 °C temperature can be reached with some scalding water.

It is intuitively clear that in any type of control system a reaction time greater than one period does not represent an inconvenience; on the contrary, delays are fundamental for the regulation of Y_t even when, without a reaction time of two or more periods, they could lead the system to destruction. With a reaction time of four periods, and with the data in Fig. 2.9, the shower would not produce oscillations and thus enable the objective to be achieved and stably maintained.

2.16.6   Direct and Inverse Control

I shall define *direct* control as the control carried out by a control system according to the [**s-o-s**] logic shown in Fig. 2.2. Many control systems operate in an *inverse* manner according to the sequence [**o-o-o**]. A simple inverse control system is the cooling system in our refrigerator. The time the compressor is turned on, X, influences in direction (**o**) the temperature of the refrigerating cell, Y, in order to adjust this in line with the temperature objective (Y^*) we have set with the memory apparatus (dial with numbered notches).

In control systems for a supermarket trolley, a hang glider, a bicycle, or a sailboat with tiller, the relations between the shifting of the angle of the bar and the angle of direction of the trolley, hang glider, bicycle, or boat are in direction (**o**). Figure 2.24 shows the inverse control system of a small motor boat with a tiller.

Increase in temperature C°/t					4	$Y_{t=0}$ =	40
Tolerance precision in degres C°					1.5	Y^* =	25
Δ Outside temperature					0.5		

t	action rate degrees/time period	Y = inside temperature	Y^* = OBJECTIVE	Variation in objective	D = disturbance	Error = Y^* - Y
0	0.0	40.5	25.0	0	0.5	-15.5
1	0.0	41.0	25.0	0	0.50	-16.0
2	0.0	41.5	25.0	0	0.50	-16.5
3	-4.0	38.0	25.0	0	0.50	-13.0
4	-4.0	34.5	25.0	0	0.50	-9.5
5	-4.0	31.0	25.0	0	0.50	-6.0
6	-4.0	27.5	25.0	0	0.50	-2.5
7	-4.0	24.0	25.0	0	0.50	1.0
8	-4.0	20.5	25.0	0	0.50	4.5
9	-2.5	18.5	25.0	0	0.50	6.5
10	0.0	19.0	25.0	0	0.50	6.0
11	4.0	23.5	20.0	-5	0.50	-3.5
12	4.0	28.0	20.0	0	0.50	-8.0
13	4.0	32.5	20.0	0	0.50	-12.5
14	-3.5	29.5	20.0	0	0.50	-9.5
15	-4.0	26.0	20.0	0	0.50	-6.0

Fig. 2.25 Simulation of an automatic on–off air conditioner

2.16.7   Simulation of an On–Off System: The Hot and Cold Air Conditioner

Figure 2.25 presents a basic numerical simulation of a hot and cold air conditioner viewed as an on–off system.
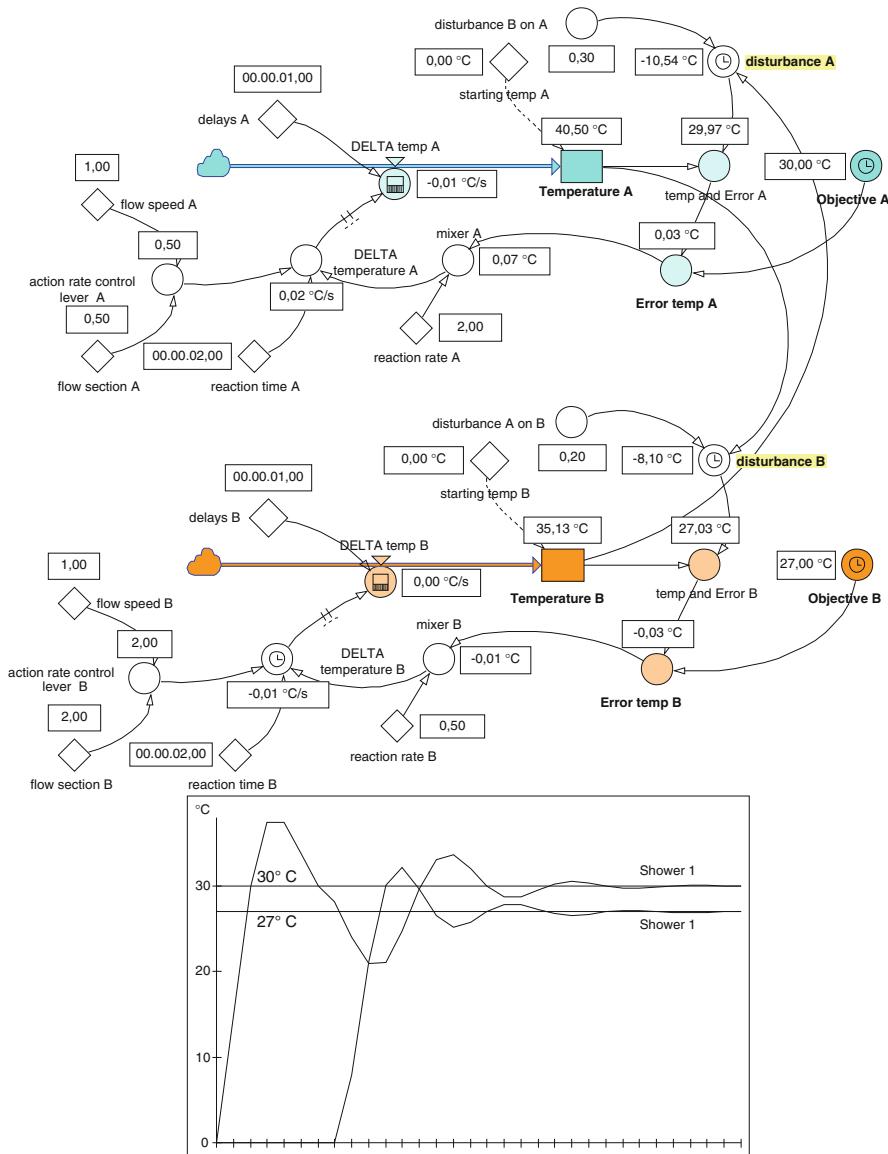
Fig. 2.26 Simulation using Powersim of two interfering showers (reaction time equal to 2 and delay equal to 1)

2.16.8   Simulation of Two Interfering Showers Using Powersim

The model of a control system for two interfering showers presented in Fig. 2.13 can be translated using Powersim into the SFD in Fig. 2.26.
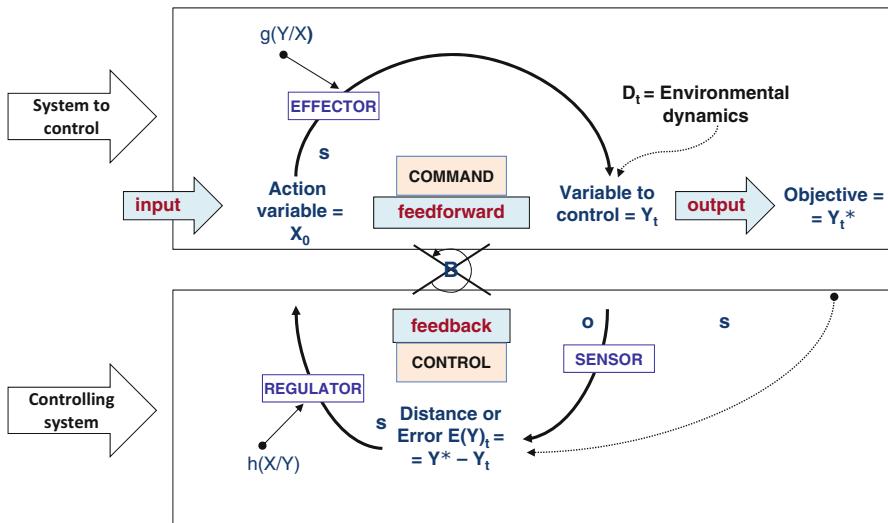
Fig. 2.27 Feedforward control system

2.16.9   *Feedforward Control*

The control systems described so far—and which are the subject of this book—act according to a *closed-loop* or *feedback* logic, in that X_{t+1} is recalculated based on the error $E(Y)_t$ and influences Y_{t+1} so as to gradually reduce the error. Nevertheless, it is possible to also plan *open-loop*, *feedforward*, or *command control systems* where, based on a *well-known model* regulating the causal relations between X and Y, the manager calculates a one-time initial value, X_0, for the control lever to trigger the dynamics of Y_t in order to achieve the objective Y^*. Once the system starts up, it is not possible to adjust Y_t, since the manager cannot modify the control lever's initial value. For this reason *feedforward* systems have no need to calculate the error in order to achieve the objective but only to "command" X appropriately in order to eliminate beforehand the error, which is already known or predictable.

Formally speaking, the *feedforward* control system has a causal *chain*, non-repetitive structure that tries to achieve the objective Y^* by preemptively setting X_0. In fact, such a system is missing an internal *controlling system* capable of intervening on the trajectory of Y_t, while this is being produced; it can only "command" the control lever X_0 in order to program the trajectory of Y_t toward Y^*, which explains why it is also called a *system of command* or *commands* (Fig. 2.27).

The feedback control systems in Fig. 2.10 can also be called *decision and control systems*, since the control of Y_t is achieved through decisions to vary X_t over a succession of repetitions of the control cycle. Feedforward systems can also be called *decision* or *one-shot control systems* since they try to reach the objective

through a single cycle and thus through a single *initial decision*. They are typically systems that "fire a single shot" to achieve the objective. The decisions are made based on precise calculations; however, when an error is detected in the attempt to achieve $Y*$ the manager cannot correct this through other regulating decisions to eliminate the error.

The archer takes account of the features of the bow and arrow, also estimating the distance, wind, and other disturbances, before precisely aiming his arrow, transmitting a direction, X_t, deemed useful for hitting the target at point $Y*$. He then releases the arrow and observes its flight. He cannot do anything further since, once the arrow has left (X_0), the trajectory Y_t is predetermined and no intervention is possible to eliminate any possible error $E(Y)$. The same is true for the hunter who fires his rifle, the space center that launches an engine-less probe, the worker who ignites a blast furnace in which he places what he deems to be the proper mixture to produce an alloy, the firm launching an ad campaign in newspapers, the cook placing what is believed to be the right amount of rice into boiling water to feed his dining companions, the farmer spreading what he considers the right amount of fertilizer for a bountiful harvest.

We can recognize *feedforward* control systems all around us, since they derive from man's innate need to achieve objectives even without being able to activate a system that can detect and eliminate the errors. Equally evident is the fact man tends to replace feedforward control with feedback control whenever possible, especially when the costs of failure in achieving the objective are very high.

The advantages of feedback control systems can be summarized in the table of Fig. 2.28.

2.16.10   The Engineering Definition of Control Systems

Figure 2.27 allows us also to point out that, according to engineering theory about control systems, the effector is considered to be a *system to control* and the detector (sensor), together with the regulator, a control system that, connected by feedback to the system to control, transforms the latter into a *system under control*, as the following precise definition explains:

> The Control System is that means by which any quantity of interest in a machine, mechanism or other equipment is maintained or altered in accordance with a desired manner. Consider, for example, the driving system of an automobile. The speed of the automobile is a function of the position of its accelerator. The desired speed can be maintained (or desired change in speed can be achieved) by controlling pressure on the accelerator pedal. The automobile driving system (accelerator, carburetor and engine-vehicle) constitutes a Control System. For the automobile driving system the input (command) signal is the force on the accelerator pedal which, through linkages, causes the carburetor valve to open (close) so as to increase or decrease fuel (liquid form) flow to the engine, bringing the engine-vehicle speed (controlled variable) to the desired value (Nagrath and Gopal 2008, p. 2)

> [. . .] The general block diagram of an automatic Control System, which is characterized by a feedback loop, is shown in Figure 1.4. An error detector compares a signal obtained through

No.	Open Loop	Closed Loop
1	Any change in output has no effect on the input; i.e., feedback does not exists.	Changes in output, affect the input which is possible by use of feedback.
2	Output measurement is not requirad for operation of system.	Output measurment is necessary.
3	Feedback element is absent.	Feedback element is present.
4	Error detector is absent.	Error detector is necessary.
5	It is inaccurate and unreliable.	Highly accurate and reliable.
6.	Highly sensitive to the disturbances.	Less sensitive to the disturbances.
7	Highly sensitive to the environmental changes.	Less sensitive to the environmental changes.
8	Bandwidth is small.	Bandwidth is large.
9	Simple to construct and cheap.	Complicated to design and hence costly.
10	Generaly are stable in nature.	Stability is the major consideration while designing.
11	Highty affected by nonlinearities.	Reduced effecy of nonlinearities.

Fig. 2.28 Comparison of open loop and closed loop control system (*source*: Bakshi and Bakshi 2009, p. 16)

feedback elements, which is a function of the output response, with the reference input. Any difference between these two signals constitutes an error or actuating signal, which actuates the control elements. The control elements in turn alter the conditions in the plant (controlled member) in such a manner as to reduce the original error (*ibidem*, p. 5).

The figure (1.4) mentioned by Nagrath and Gopal is reproduced in Fig. 2.29; as we can see, the feedback control system perfectly corresponds to the one in Fig. 2.10, which can be arrived at through simple position changes of the various elements.

Figure 2.30 illustrates a more analytical model which perfectly corresponds to that in Fig. 2.29 (taken from Varmah 2010), which also highlights the input transducer, that is, the apparatus that transforms the command input (for example, the indication of the desired temperature on a graduated scale) into the reference input, that is, the objective in terms that can be read from the control system. These typically engineering apparatuses enrich the technical model of a control system in Fig. 2.10 while not reducing its generality.

In this book I have adopted the *language* of systems thinking, feeling it appropriate to focus on the controlled variable, Y_t, and to assume that, if its dynamics lead
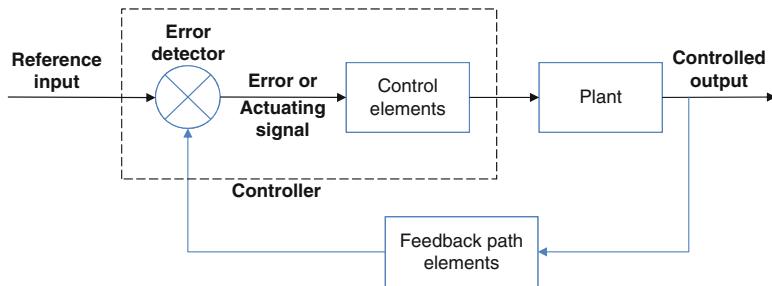
Fig. 2.29 Model of feedback control systems from an engineering perspective (*source*: Nagrath and Gopal 2008, "Fig. 1.4. General block diagram of an automatic Control System," p. 5)
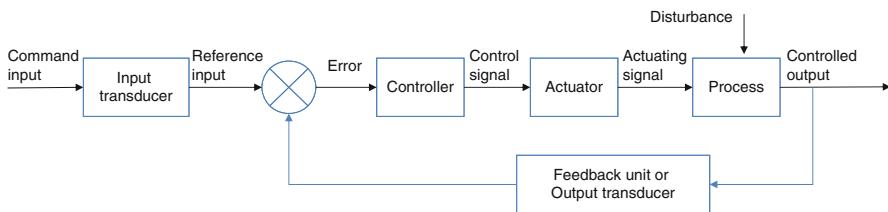
Fig. 2.30 Engineering model of feedback control systems with the essential terminology (*source*: Varmah 2010, "Fig. 1.2 Closed-loop Control System," p. 4)

it to converge to a value Y^*, defined as a constraint or objective, then its dynamics are governed by a control system. Thus, the term "control system" indicates the entire logical *loop* between the controlled variable Y_t and the control levers X_t, through the error $E(Y)$, independently of the physical, biological, or social structure of the mechanisms that make Y_t tend toward Y^*.

2.16.11   An Analytical Appendix: The Optimal Value of the Lever and the Optimal Control Period in Non-symmetrical Systems

Section 2.3 presented the general model of a control system in the form of recursive equations, as these can be immediately grasped and easily applied using both spreadsheets and general programming language. Nevertheless, it is appropriate, at least for the sake of thoroughness, to provide a brief analytical discussion that can also help in understanding the operational logic of simulation programs for stocks and flows. *In any event*, the concepts in this section are not indispensable, *and the reader can even go back to them later*.

I shall start with the simplest model (the mono-lever one) of a symmetrical control system ($h = 1/g$) in which the action equation, which enables us to calculate the optimal value Y^*, has the following simple form:

$$Y^* = g\, X_{n*} + D \qquad (2.14)$$

where X_{n*} represents the optimal value of the control lever and D the sum of the external disturbances. Let us assume the initial value of the system is:

$$Y_0 = g\, X_0 \qquad (2.15)$$

Subtracting (2.15) from (2.14) we obtain:

$$Y^* - Y_0 = g\, X_{n*} - g\, X_0 + D$$

which immediately allows us to determine the value of X_{n*}, since $E = Y^* - Y_0$:

$$\frac{Y^* - Y_0 - D}{g} = X_{n*} - X_0 \qquad (2.16)$$

Example 1: the objective is $Y^* = 40$, with $Y_0 = 0$, $g = 2$, $h = 1/2$, and $D = 0$; from (2.16) we immediately obtain: $X_{n*} = 20$, since $X_0 = 0$.

Example 2: the objective is $Y^* = 80$, with $g = 6$, $h = 1/6$, and $D = 20$; referring to (2.16) we get $X^* = 10$.

Example 3: the same values as in *Example* 2, but with $D = -20$; from (2.16) we obtain $X^* = 16.67$ as the optimal value of the lever, as the reader can verify.

Equation (2.16) clearly shows that the reaction rate, "h," is not an essential datum in the model in Fig. 2.2, since it corresponds to the reciprocal of "g"; this relationship creates a system of symmetrical control. Nevertheless, its introduction in models is useful for giving a logical completeness to the loop in Fig. 2.2, since it enables the optimal value of the control lever to be calculated immediately in one sole step.

In order to identify a general solution for calculating the optimal value, X_{n*}, for the control lever in the case of an *asymmetrical* system, I propose this simple line of reasoning.

Let us go back to (2.16) and, to simplify things, assume that $X_0 = 0$, $Y_0 = 0$, and $D = 0$, and that $h° \neq (1/g)$; let us also assume that $g° = (1/h°)$ is the action rate corresponding to $h°$, so that we can write the control system, in its minimal asymmetrical form, as follows (omitting time references):

$$\begin{cases} g\, X = Y \\ E + Y^* - Y \\ X_n = (1/g°)E \end{cases} \qquad (2.17)$$
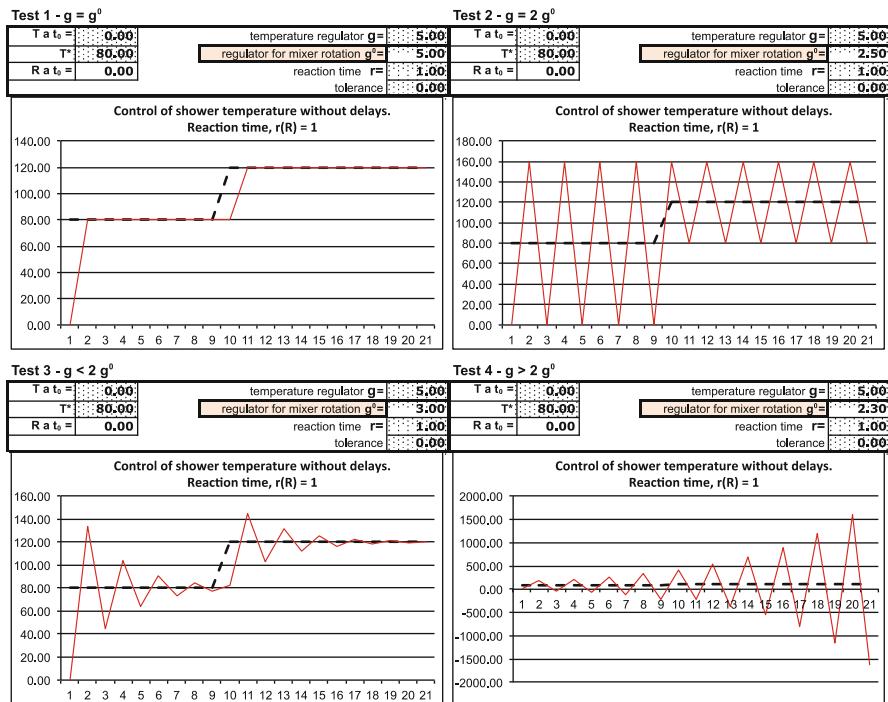
Fig. 2.31 Dynamics of non-symmetrical temperature control systems

Adopting the value for the lever X_n calculated in the third equation, if this value is optimal, we should obtain the objective value, Y^*, calculated in the first equation in (2.17):

$$g\frac{1}{g^\circ}E = Y^* \tag{2.18}$$

However, (2.18) shows that the lack of symmetry affects the dynamics of the system when the initial value of X is different from the optimal one:

– If $g < g^\circ$, then the system underestimates X_n and does not achieve Y^*, thereby producing a new error to be eliminated through a new calculation of X_n; the system achieves its objective through a succession of iterations.
– If $g > g^\circ$, the system overestimates the value of the lever and exceeds the objective; depending on the value of g°, the system may converge or diverge.

Figure 2.31, which is derived from Fig. 2.8, represents the results from four tests carried out on the temperature system, without delays and with a variation in the objective at $t = 9$. In general:

– Test 1—for $g = g^\circ$ the system is perfectly symmetrical.
– Test 2—for $g = 2g^\circ$ the system oscillates over constant periods.

– Test 3—for $g < 2g°$ the system converges with diminishing oscillations.
– Test 4—for $g > 2g°$ the system diverges.

A final observation. Considering—for simplicity's sake—only symmetrical systems, (2.16) not only allows us to quantify the optimal value of X in *continuous phase* control systems—as the previous examples have shown—but is also equally useful for *discrete step* control systems, since it permits us to determine the *control period*, that is, the number of periods at the end of which Y attains $Y*$. Let us assume $x*$ to be the rate of variation of X; from (2.16) above, $X_{n*} = n*x*$, with $n*$ being the number of times step $x*$ must be repeated in order to obtain the optimal value of the lever X, which gives us the value $Y*$. The number $n*$, which represents the number of periods the step must be repeated, expresses the control period needed to bring Y to the objective value $Y*$.

From (2.16) we automatically obtain "$n*$":

$$\frac{Y* - Y_0 - D}{g} + X_0 = n*x* \tag{2.19}$$

Example 4: the objective is $Y* = 80$, with $g = 5$, $h = 1/5$, and $D = -20$, assuming a fixed step, $x* = 1$; we immediately get $n* = 20$, which reduces to 10 when $x* = 2$.

Example 5: the objective is $Y* = 120$, with $g = 10$, $h = 1/10$, and $D = 20$, assuming a fixed step, $x* = 2$; we immediately get $n* = 5$.

Equation (2.19) allows us to make another simple, though important, calculation. Let us assume we have set the condition that the control must be carried out during a predetermined period, $n*$, which represents a second time objective. It is clear that in order to simultaneously achieve the objectives $Y*$ and $n*$, the manager must be able to set the *action rate*, g (the capacity), for the lever X.

Thus, (2.19) allows us to calculate the action rate, $g*$, that enables Y to attain the value $Y*$ during a period of $n*$ instants. Setting $X_0 = 0$, we immediately get:

$$\frac{Y* - Y_0 - D}{n*x*} = g* \tag{2.20}$$

Example 6: the objective is $Y* = 120$, $D = 20$, $x* = 2$, and $n* = 5$, we immediately get $g* = 10$, which decreases to $g* = 4$ if the step increases to $x* = 5$.

2.17   Summary

We have learned that:

(a) The *logical structure* of single-lever control systems (Sect. 2.2) always includes the following elements (Fig. 2.2): X_t; Y_t; $Y*$; $E_t = Y* - Y$; D_t; $g(Y/X)$; $h(X/Y)$; $r(X/Y)$ and delays.

(b) The dynamics of the variables X_t and Y_t can be determined with the recursive equations, without delay, indicated in Sect. 2.3; or with delay, as indicated in Sect. 2.8.

(c) Every control system has its own *logical* and *technical structure* (Sect. 2.9 and Fig. 2.10).

(d) The *technical structure*—or *chain of control*—is a "machine" composed of mechanisms (or apparatuses) that develop the processes that produce the variations in the variables *X*, *Y*, and *E*; it includes: the *effector*, the *sensor*, the *regulator*, the *information network*.

(e) Single-lever control systems can interfere or be linked (Sect. 2.14) in order to form a wider system; if there are delays, the global system can produce oscillating dynamics (Figs. 2.11, 2.12, and 2.13); nevertheless, the presence of a reaction time of 2, 3 or even more time units makes it possible to achieve *Y**.

(f) In any control system we can find a *manager* and a *governor* (Sect. 2.10).

(g) A control system is developed by the *designer* and built by the *constructor* (Sect. 2.11).

(h) The general model of control system can represent all situations of change (Sect. 2.15), in particular:

  – The problem-solving process (Fig. 2.14)
  – The decision-making process (Fig. 2.15)
  – The change-management process (Fig. 2.16)
  – The performance control systems (Fig. 2.17)

Chapter 3
The Ring Variety: A Basic Typology

It seems that cybernetics is many different things to many different people, but this is because of the richness of its conceptual base. And this is, I believe, very good, otherwise cybernetics would become a somewhat boring exercise. However, all of those perspectives arise from one central theme, and that is that of circularity [. . .] While this was going on, something strange evolved among the philosophers, the epistemologists and the theoreticians: they began to see themselves included in a larger circularity, maybe within the circularity of their family, or that of their society and culture, or being included in a circularity of even cosmic proportions. What appears to us today most natural to see and to think, was then not only hard to see, it was even not allowed to think! (Heinz Von Foerster 1990, online).

Chapter 2 presented the logic of control systems and described the technical structure that produces the general logical structure. I have thought it useful to begin with the simplest control systems, single-objective (a single Y and single Y^*) and single-lever (a single X) ones, in order to present and simulate several elementary systems drawn from our daily experiences, in which the action function determining the dynamics of the variables (g and h) are represented as simple constants. This chapter presents several fundamental classes of control systems, distinguished by the type of human intervention in the control or by the nature of the objective: natural and artificial systems, manual or automatic cybernetic systems, quantitative and qualitative systems, attainment and recognition systems, steering and halt control systems, fixed- and variable-objective systems (or systems of pursuit), tendential and combinatory systems, as well as systems distinguished by other features. Finally, we shall consider various forms of interconnection among control systems and the holarchies of control systems.

3.1 Manual and Automatic Control Systems: Cybernetic Systems

The first useful distinction to make—of an engineering nature—is that between *manually controlled systems* and *automatic control systems*. This distinction is based on the presence or lack thereof of human intervention in the chain of control

> Systems [which] involve continuous manual control by a human operator . . . are classified as manually controlled systems. In many complex and fast-acting systems, the presence of the human element in the control loop is undesirable because the system response may be too rapid for an operator or the demand on the operator's skill may be unreasonably high. . . . Even in situations where manual control could be possible, an economic case can often be made for reducing human supervision. Thus in most situations the use of some equipment which performs the same intended function as a continuously employed human operator is preferred. A system incorporating such equipment is known as an automatic Control System (Nagrath and Gopal 2008, pp. 3–4).

> In short, a Control System is, in the broadest sense, an interconnection of the physical components to provide a desired function, involving some kind of controlling action in the system (Bakshi and Bakshi 2009, pp. 1–3).

The above-mentioned distinction is appropriate but, precisely because it derives from an engineering view of control systems, which focuses on the control of systems understood as mechanical apparatuses, we need to make several observations and additions. Above all, recall that the Preface of this book presented several examples that show that, in terms of quantity and quality, most of the *automatic control systems* are *natural*, biological, or even social systems where human intervention is not possible. In fact, all biological systems that control the existence of all living beings, from the tiniest cell to organizations and society, are natural systems. Such systems act on organs of all types and varieties that usually cannot be directly controlled (breathing, heartbeat, immune system, etc.), on the body's structural components, even on the biological clocks that regulate most of our biological cycles, etc. Control systems that operate in the physical and meteorological world to produce clouds or dew or in ecological systems to restore balance in prey, and predator populations also represent natural systems.

In his daily activities man represents a formidable control system that produces hundreds of control activities each minute. None of us bump into other passers-by when we encounter them or cross the street without first observing the traffic light or how far away the moving cars are. We check our position and that of our limbs and set objectives for our behavior. We drive our car to the desired destination and in doing so we grab and maneuver the driving wheel, reach and operate the various levers, adjust from time to time the direction of the car, the speed and the distance from other cars, etc. I shall consider these systems in more detail in the second part of the book.

All control systems designed and constructed by man, whether or not *manually controlled* or *automatic* control systems, are *manmade control systems*. The distinction between *manual* or *automatic* systems is not simple to make, since human intervention in *artificial* control systems can take quite different forms and intensities and act differently on the "apparatuses" that make up the chain of control.

It can be useful to refer to Fig. 3.1, which is directly derived from Fig. 2.10. We can define as *automatic* those control systems in which human intervention is limited to externally setting the objectives, Y^*, which the variable to control, Y, must achieve (point ①). The objective represents the only "information" for the system regarding the dynamics to set for Y; *sensor*, *regulator*, and *effector* are natural or artificial apparatuses on which man does not intervene, unless possibly to supply energy and for maintenance. Thus man assumes the role of *governor* of the
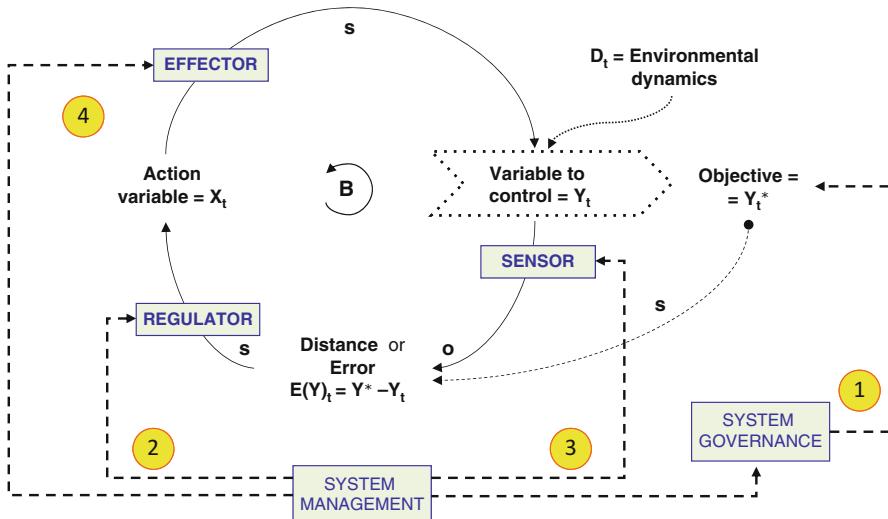
Fig. 3.1 Types of human intervention in artificial control systems

control system; by setting the latter's objectives, he guides the system in the action space of Y but does not take on the role of *manager*, since management is an integral part of the regulatory apparatus. We can define as *manually guided* control systems those where man becomes the manager of the system and intervenes through his own decisions and actions on the regulator in order to determine the values of the control lever X (point ②). The objectives can be assigned by the manager himself or by an outside governor.

An *automatic* control system guided toward a fixed objective determined by an outside *governor*, but guided by a *manager within the chain of* control, is also defined as a *cybernetic system*; the *cybernetic system* appears to be closed within itself, and the only link with the environment is represented by the specification of the objective Y* by the governor.

The definition of *cybernetic system* can be linked to the idea of "guiding" a system in time/space (for example, a ship); that is, of controlling its movement toward an objective (port of arrival) by means of an automatic control system in which the human manager (helmsman) is replaced by a mechanical manager (GPS-guided rudder), thereby creating an automatic *Ring*. The science that studies how to produce automatic systems was founded by Norbert Wiener, who defined this as the science of the control (and guidance) of systems. In fact, Wiener's book is entitled: *Cybernetics, or control and communication in the animal and the machine.*

> We have decided to call the entire field of control and communication theory, whether in the machine or the animal, by the name *Cybernetics*, which we form from the Greek χυβερνητητ [sic] or *steersman*. In choosing this term we wish to recognize that the first significant paper on feedback mechanisms is an article on governors, which was published by Clerk Maxwell

in 1868, and that *governor* is derived from a Latin corruption of χυβερνητητ [sic]. We also wish to refer to the fact that the steering engines of a ship are indeed one of the earliest and best-developed forms of feedback mechanisms (Wiener 1961, p. 11).

Norbert Wiener—precisely because he viewed cybernetics as the *theory of the guidance* or *governance* of systems; that is, the study of the *logic* according to which the helmsman (manager) guides the ship to the port (objective $Y*$) indicated by the ship owner—assigned a fundamental role to both the *communication* of the objective to the manager and the latter's orders (control) to modify the route (Y) by means of the rudder (X). For cybernetics, the true input of the control system is the information on the amount of error ($E(Y)$) with respect to the objective, so that the manager can rectify the value of the control lever to guide the "ship" to port. In another important book, *Cybernetics and Society*, Norbert Wiener wrote:

> In giving the definition of Cybernetics in the original book [Cybernetics], I classed communication and control together. Why did I do this? [...] When I communicate with another person, I impart a message to him [...] Furthermore, if my control is to be effective I must take cognizance of any message from him which may indicate that the order is understood and has been obeyed. [...] When I give an order to a machine, the situation is not essentially different from that which arises when I give an order to a person. [...] Naturally there are detailed differences in messages and in problems of control, not only between a living organism and a machine, but within each narrower class of beings (Wiener 1954).

Precisely to the extent it is a science of guidance and communication, cybernetics today is mainly associated with the control of machines and mechanical systems. None other than Ludwig von Bertalanffy, considered to be the founder of General Systems Theory, recognized that:

> [A] great variety of systems in technology and in living nature follow the feedback scheme, and it is well-known that a new discipline, called Cybernetics, was introduced by Norbert Wiener to deal with this phenomena. The theory tries to show that mechanisms of a feedback nature are the bases of teleological or purposeful behaviour in man-made machines as well as living organisms and in social systems (von Bertalanffy 1968, p. 44).

Nevertheless, as Ross Ashby, another founder of cybernetics, clearly states that what cybernetics examines are not machines but the control systems that allow these to function in order to produce the desired dynamics.

> Many a book has borne the title "Theory of Machines", but it usually contains information about *mechanical* things, about levers and cogs. Cybernetics, too, is a "theory of machines", but it treats, not things but *ways of behaving*. It does not ask "what *is* this thing?" but "*what does it do?*" [...] It is thus essentially functional and behaviouristic. [...] Cybernetics deals with all forms of behaviour in so far as they are regular, or determinate, or reproducible. The materiality is irrelevant, and so is the holding or not of the ordinary laws of physics. [...] The truths of cybernetics are not conditional on their being derived from some other branch of science. Cybernetics has its own foundations. [..] Cybernetics stands to the real machine—electronic, mechanical, neural, or economic—much as geometry stands to a real object in our terrestrial space (Ashby 1957, pp. 1–2).

Even more recent authors have recognized that cybernetics does not have a single, so-to-speak practical, function but represents instead a general *discipline* for observing and interpreting the world:

> Cybernetics is not exactly a discipline. Rather, it should be thought of as an approach which helps thinkers and practitioners across a whole range of different disciplines secure a vantage from which they are likely to gain a more complete understanding of their own area of practice. Its focus is upon meaningful and effective action and behaviour in whatever domain of human activity (Velentzas and Broni 2011, p. 737).

I have provided the above quotes because they show that even cybernetics as a *discipline* studies the *logical models* of control systems not necessarily their *technical realization*. I would like to specify that this book is not a cybernetics text, since it does not investigate how control is effected in systems containing an objective – machines, organisms, societies, organizations, etc. – but how such systems can be interpreted as Control Systems. I shall not attempt to clarify how the eye focuses on objects but how the eye is "in an of itself" the Control System that focuses on objects. Nor shall we investigate how firms can be controlled but how they are Control Systems "in an of themselves".

I wish to make also clear that this book considers the general theory of control systems in a different, probably broader, way than cybernetics; I shall not consider only automatic–cybernetic systems but *natural systems* as well, where human intervention is absent, and *manual systems*, where human intervention not only is connected to the communication of orders and the reception of answers but also can directly intervene on the chain of control itself. I would note that *natural systems* are not only cybernetic. A control system can be defined as *natural* if the following twofold condition applies: (1) it is *automatic* and (2) it has no *outside governor*, since the objectives are in the form of constraints, limits, or endogenous conditions. The eye is a natural control system, since it tries to control the quantity of light that must reach the retina (constraint Y^*) through the automatic control system of the pupil, which, through small, speedily contracting muscles, rapidly varies its diameter. All (human) eyes function in this manner.

The general theory of control systems does not consider only the logic behind the functioning of control systems—also proposing a significant typology of such systems—but becomes a true control discipline whose objective is to demonstrate how fundamental control systems are for the existence of our planet, aiding the reader to identify which control systems are operating in each sector of our knowledge. As a discipline it also seeks to study how every *automatic* control system can grow and become *manually* guided, and how every *manually* guided system can transform itself into an automatic system in which management can become part of the control apparatus itself.

I would also note that the *natural* control systems for our physiological control are not only automatic; we can become the outside managers of the chain of control. If the light is too strong we lower our eyelids, wear sunglasses, or seek out the shade. If it is hot, the natural body temperature control system activates the sweating process; however, if the heat is still too high and sweating does not sufficiently lower our body temperature, we take off some of our clothes. If we still feel hot we take a cool shower; and if the heat persists for a long time we activate the air conditioner or take a holiday in the mountains.

Going back to Fig. 3.1, we can finally observe that man can direct the control process by intervening in various ways on the technical structure of control systems; (a) he can function as a *sensor* (point ③) by detecting the errors directly

with his own organs: by feeling when the water or air temperature is too hot; by perceiving a too high or too low sound volume; by noticing that the automobile is veering dangerously toward the guard rail; by tasting the risotto to see if it is done, etc.; (b) he can function as the *effector* (point ④) by stepping on the brake pedal, turning the handle bars on the bicycle, opening/closing a valve, etc.; (c) he can be the regulator (point ②) that, based on his knowledge, decides when or to what extent to maneuver the control lever; (d) finally, he can intervene globally (points ① + ② + ③ + ④) by forming the entire chain of control, where man becomes "the" control system. To convince yourselves of this, just "look around."

Finally, the control discipline reveals the importance of the *observer's point of view* in recognizing and constructing models and simulating the behavior of control systems. As taxi passengers we view ourselves as the *governor* of the transportation system (we give the destination) and observe the car as a control system (machine) whose manager-controller is the taxi driver; if we are driving our car we believe we ourselves are the control system for speed and direction, which we perceive with our own senses and modify with our own limbs, considering the car as the effector toward the desired destination; if an alien intelligence observes the Earth from a UFO several miles above, and sees a car winding through a network of roads, it would probably think it was seeing natural control systems operating, changing direction and speed without any objective or human intervention, without ever colliding or veering dangerously away from the flow of traffic, exactly as we imagine the ants do that are marching along the plants in our garden.

3.2 Quantitative and Qualitative Control Systems: Attainment and Recognition Control Systems

The general model of a control system shown in Fig. 3.1 can be applied to both *quantitative* and *qualitative* control. The control is *quantitative* if the variable Y_t represents a *quantitative measure* (in any of its multiform expressions) and the objective is a value that Y_t must reach (*set reference*) or a set of values to be achieved (*track reference*).

I shall define *attainment control systems* as those control systems that act to "attain" a *quantitative* objective, independently of the control context. All the control systems models presented so far—audio, temperature, directional, speed control, etc.—are cases of *attainment control systems*. A simple observation will immediately reveal that *attainment control systems* are all around us, since we encounter them whenever both the variable Y_t and the objective Y^* can be quantified or are separated by a distance that is in some way measurable. Figure 3.2 illustrates the typical example of the control system we activate when we want to grab something with our hand; for example, the gesture of grabbing a cigar or a pencil as described by Norbert Wiener:

> If I pick up my cigar, I do not will to move any specific muscles. Indeed in many cases, I do not know what those muscles are. What I do is to turn into action a certain feedback mechanism; namely, a reflex in which the amount by which I have failed to pick up the cigar is turned into a new and increased order to the logging muscles, whichever they may be (Wiener 1954, p. 26).
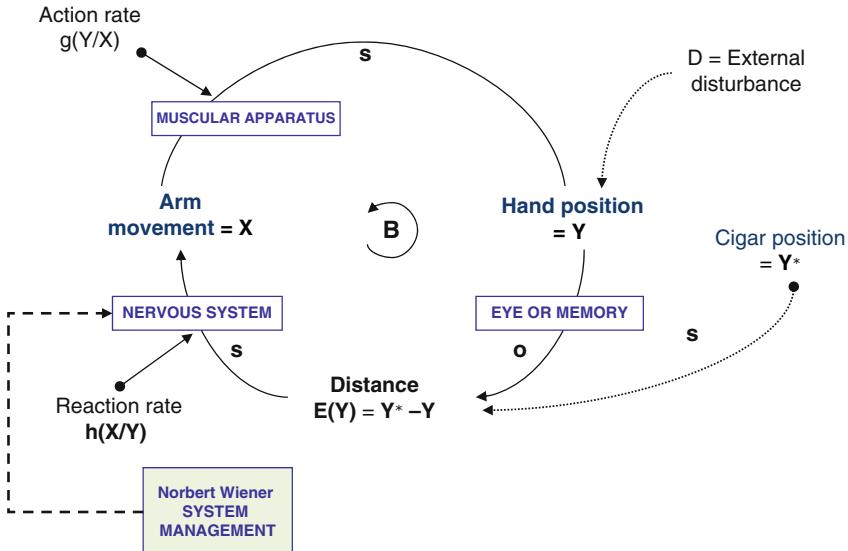
Fig. 3.2   Attainment control system. Norbert Wiener picks up his cigar

Other less obvious examples of such systems are those structured to control that given levels of average hourly productivity in a factory are achieved, that a desired monthly level of *operating cash flow* in a certain department is achieved, that the public debt does not exceed $2 trillion, that a given threshold of dust particles in the air is not exceeded, that our body temperature does not go beyond its normal limits, that microwave ovens operate for the desired length of time, or that our eye or our reflex camera can bring images into focus, etc.

We can distinguish among the *attainment* control systems based on the nature of the objective:

1. *Goal-seeking systems*: those in which Y* represents a quantitative *objective*: achieve a result, or attain or maintain operating standards;
2. *Constraint-keeping systems*: in which Y* represents a *constraint* to respect or a *limit* not to be exceeded; these are usually *halt systems* (see Sect. 3.3 below).

Recalling the concept of a control problem (Sect. 2.1), we can further distinguish among control systems as follows:

1. *Regulator systems*, where Y* (objective or constraint) is a specific value (goal, constraint, or limit) that Y_t must reach and maintain over time;
2. *Tracking systems*, or *path systems*, where Y* (goal, constraint, or limit) is a *trajectory*, that is, a sequence of values, Y_t*, however formed, which Y_t must follow.

Despite the abundance and ease of observation of quantitative control systems, we must nevertheless be aware that there exists an equally broad and important class of

*qualitative control system*s in which both the variable Y_t and the objective $Y*$ are qualitative in the broadest sense of the term (objects, colors, flavors, forms, etc.). If we want a white wash load ($Y*$, qualitative objective) we must modify the state of cleanliness (Y_t, qualitative variable) until we are satisfied ($E(Y)$ reduced to zero). If we want a perfectly cooked roast ($Y*$, qualitative objective) we must modify the cooking state (Y_t, qualitative variable) until we are satisfied ($E(Y)$ reduced to zero).

It is immediately clear that in qualitative control systems the manager–governor faces the problem of coherently expressing the qualitative variables based on a scale of differentiated states so that the detector apparatus can compare the present state, Y_t, and that representing the objective, $Y*$, in order to determine the variance. How do we determine the state of "cleanliness" of laundry? How can we detect the cooking state of food? The problem of how to graduate the "states of variety" in the qualitative dimensions is often left to the experience of the manager. The cook knows if the food is well-flavored, not salty enough or almost done; every housewife knows when the laundry or dishes are cleaned to the right extent. Nevertheless, procedures to determine the "states of variety" of the qualitative variables have been worked out, the most important of which are presented in Sect. 3.9.9. However, these examples are not enough to make us aware of the generality of *qualitative control system*s; in order to fully understand their *modus operandi* and to introduce the very relevant class of *control systems of recognition*, I now propose four other examples.

Let us first imagine we have to buy a button to replace the one we lost on our grey suit ($Y*$). We go to the store that sells buttons to buy the one we need and are immediately embarrassed because the shop owner asks us the button size. We remember it is "around" 2 cm in diameter. Then, when we are asked to specify the color, we panic because the owner presents us with dozens of buttons (Y_n) of that diameter in various shades of grey, and we do not know how to *recognize* (identify) the one closest to the desired shade, assuming the size is correct. There are four possibilities for arriving at the optimal choice; the first is to have at hand the grey suit with the lost button, so that we can compare the remaining buttons ($Y*$) with those the shop owner has in stock. If we are lucky we will find a button that we *recognize* as identical to the desired one. If we do not have the suit with us we can, as an alternative, at least have a photograph of the button to replace ($Y*$). We have had the foresight and wisdom to have taken a photo of the suit and its buttons on our SmartPhone before leaving home, which we now show to the shop owner. In this case as well we can easily *recognize* the button to purchase, thereby solving our problem. As a third alternative, before leaving home we could have compared our suit buttons with a color chart (catalogue) containing a wide variety of shades of grey, numbered progressively (or by name), and identified the number that corresponds to the shade of the suit buttons. Here, too, if the shop owner has the same chart we can precisely *recognize* the button. It is perfectly clear that if we cannot compare the buttons on the shop counter with an identical copy of the original one, using a graphical model or the shades of grey on a chart, the only chance left to recognize the desired button is to hope in our prodigious visual memory, together with a good deal of luck, in order to choose without having the actual sample at hand.

Let us now consider a second example. Our university mate, George, who I have not seen for years, telephones to invite me to a party he will be attending. I am happy for the chance to get together with my old friend. I arrive at the party only to realize there are around 300 people in a large reception room; I start looking for my friend George based on my recollection of his facial features. Walking around the room I observe one person after another, obviously only the males. "This isn't George; it resembles him, but it's not him ....." After numerous attempts, I spot someone at the end of the room who *could* be my friend; however, when I approach him I realize he is not George, though he quite resembles him. I continue searching for several minutes until I see a person who, despite his hair color and some wrinkles, appears to be my friend. I have finally *recognized* George. The same recognition would probably have been made by someone to whom I had assigned the task and to whom I had given a photo of George. Of course, in both cases it would not have been necessary or appropriate to compare the fingerprints of those present at the party with those the army had of George. But in other cases, which I shall leave to the reader to imagine, fingerprints would be useful for recognition (Sect. 3.9.3).

As a third example, consider an eagle that, in its majestic flight, crosses a valley in search of food (Y^*). Observing the ground from a certain height with his sharp eyes, he notices many animals (Y_t) that would be suitable as food. He identifies a huge animal and immediately compares it with the *catalogue* of possible prey his life experience has compiled year after year in his memory; in this catalogue (memory) that animal is *identified* as a bison, and his memory reminds him that it is too large and cannot be lifted into the air; thus it does not represent a prey. He then *identifies* a snake, but it is too close to the rocks and thus difficult to capture. In the middle of the alley he observes an animal of suitable body mass, and his catalogue (memory) *identifies* it as a small deer that seems not to have noticed the approaching eagle. The eagle quickly swoops down and captures the deer and takes it to its nest.

We have taken a course for *sommelier*, tasting dozens of different wines, memorizing for each its taste, color, aroma, and other organoleptic features, resulting in our compiling a catalogue. At the final exam to earn qualification as a *Sommelier* we are given a wine (Y^*) and asked to indicate the name, year of production, and other characteristics determined by the jury. We taste the wine (Y^*) and immediately our mind scans the "catalogue" of stored "wine models" (Y_n), calling them up one by one until we can identify the wine objective as a 2009 *Brunello di Montalcino*.

On careful reflection, the *identification* process of animals carried out by the eagle, or of wines by the sommelier, is different from the *recognition* of my friend George at the party or of the right button for my grey suit. In *recognizing* George (or the button) I compare "real objects" that *scroll by* before me (Y_n), $n = 1, 2, \ldots$, with a "model-objective" (Y^*) and, when there is a correspondence, I am able to *recognize* a friend (a button). The eagle (the aspiring sommelier), on the other hand, from time to time compares the "stored models, or catalogue," Y_n, that *scrolls by* in his mind one after the other with a "real object" (Y^*); that is, with the animal it

observes (the wine he tastes). Only when it/he succeeds in *identifying* a prey (or a type of wine) is the qualitative objective (prey, wine) achieved.

I can now generalize. In all situations where we have to *recognize* some "model-objective" (George, button) by comparing it with "real objects" from a set that is sequenced (invitations to the party, buttons on the counter) we activate a *system of recognition* where the qualitative variable to control, Y_n, represents the "real objects" that we *scan* in order to recognize the "model-objective" Y^*. In situations where we instead have to *identify* a "real-object" set as an *objective* Y^* by *scanning through* a "catalogue of models," Y_n, we activate an *identification system*.

It is not always easy to distinguish between these two types of system since, despite their operational differences, systems of *recognition* and of *identification* follow the same logic and can be described in a similar way, if we consider that, in a strict sense, all observed *objects* can be considered, from a technical point of view, as models of *objects* or as *technical descriptions* of objects (Sect. 9.4).

In order to *recognize* a model-objective or *identify* a real-object-objective, Y^*, we require:

- An *archive* of "N" *objects* (buttons, party invitations) or *models* (preys, wines tasted and memorized) represented by a variable Y_n, $n = 1, 2, \ldots, n, \ldots, N$.
- An *effector* apparatus that scans the *archive* by changing the value of Y_n over a time sequence based on some predetermined criterion (alphabetical order, size, sex, location, route, etc.).
- A *sensor* apparatus to compare the value of Y_n with the object/model-objective to recognize, Y^*, in order to determine, through preestablished rules, if $Y_n = Y^*$ or if there is an error $E(Y)$.
- A *regulatory* apparatus that evaluates $E(Y)$ and decides whether or not to continue with the *scrolling* of the catalogue ($n + 1$, $n + 2$, etc.).
- An *action lever*, X_n, that orders the effector to *scroll through* the catalogue for successive "n" if $E(Y) \neq 0$ (according to the predetermined rules of comparison).

I thus propose to consider *recognition systems* as a general class of qualitative control systems and *identification systems* as a subclass. In any event, it is clear that the process of *recognition* or of *identification* just outlined corresponds perfectly with the general logic of control examined in Chap. 2 and that this process is carried out by a qualitative control system.

The model in Fig. 3.3 illustrates the logic of *systems of recognition*, which can be applied, for example, to a person, animal, or machine that explores the environment—that is, the archive of possible objects or models—in order to recognize an object or model-objective and continue the exploration until the error is eliminated.

As Fig. 3.3 shows, all *recognition control systems* must possess special *apparatuses* in their chain of control. The *sensor* organs must be guided by special *algorithms* that permit us to make a description of the observed model. The *sensors* can make a comparison with the model-objective only if they are guided by some comparison *algorithm* capable of carrying out the comparison between the observed description and the objective one. Finally, the *regulation* apparatus must
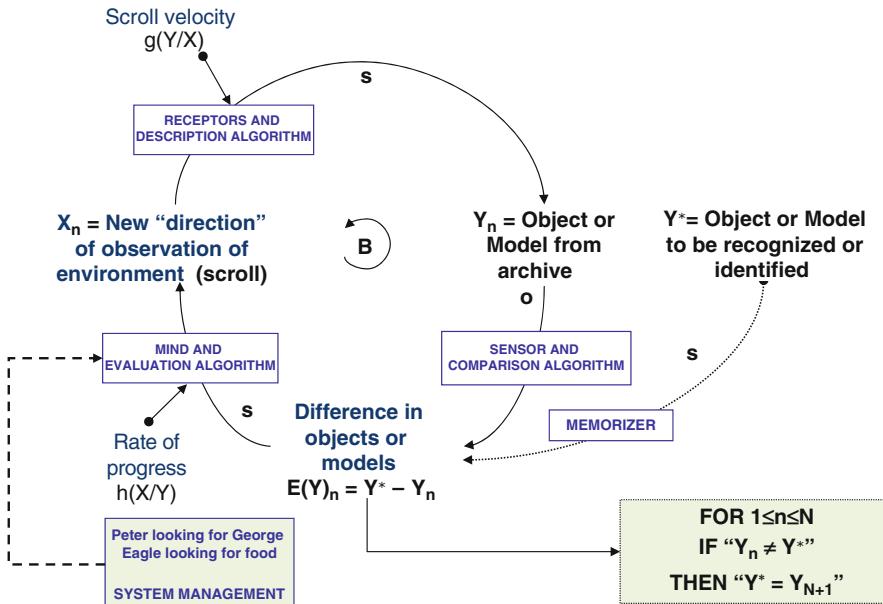
Fig. 3.3 Control system of recognition-identification

be guided by some error evaluation and decision-making algorithm in order to decide how to act on the control lever (Zhao and Chellappa 2006).

Figure 3.3 allows us to explore the infinite world of *qualitative control systems of recognition* that permeate our life as well as that of all animals equipped with sensor organs that make it possible to compare models and objects. In living systems the model-objective is memorized mentally by both the predator, which must identify a prey, and the prey, which must recognize the predator; by both the pollinating bee that must identify a flower to pollinate and the warrior who must distinguish friends from enemies; by both the infant that must recognize his parents among a multitude of people and those who must recognize the person calling on the phone. How would our life be if we could not recognize people, objects, words, etc.? Consider how many people we can recognize on a daily basis from facial features as well as from tones of voice, the way of striking keys on a keyboard or of stopping a car in front of a house, etc.; not to mention the ability to read a daily newspaper and recognize keys in a larger set, a badge, a credit card, the telephone number of friends, etc.

Many *control systems of identification* are artificial and can scan different types of "catalogue" at high speed in order to recognize the model-objective, as in the case of the facial and expression recognition systems used by the police or by intelligence agencies in airports (Sect. 3.9.3). If there is a substantial inherent difference between asking my son: "let me know when you see (recognize) mom arriving" and ordering the policeman: "check if this digital fingerprint is in

(recognized by) the AFIS," from a formal point of view these are two typical recognition systems: the former one human and the latter computerized.

Most *control systems of recognition-identification* particularly in the animal kingdom are usually linked to some *attainment control system* with a spatial objective; when the chameleon recognizes its prey and locates it at a given spatial point, it tries to *reach it* with the control system that uses an extendable tongue; when the hawk recognizes a mouse it dives to reach it; when a doctor recognizes a pathology he starts to initiate a control system to eliminate it.

The c*ontrol systems of recognition-identification* are fundamental for the survival of individuals and species. They operate not only at a macrolevel but a micro one as well, inside our bodies. Our immune system represents the most powerful *recognition-identification* system in our bodies, without which our existence would not be possible. The immune system protects the organism from aggression from pathogenic agents, recognizing in a highly specific manner millions of substances extraneous to the organism: the antigens, viruses, or bacteria. This defense is based on the action of lymphocytes that produce the antibodies that recognize the antigens; the lymphocytes then attach to the antigens in order to facilitate the work of the macrophages that then proceed to destroy them. The system acts at the microlevel through the continual production (X) of different antibodies (Y_n) until one of these couples to (recognizes) the intruder (the antigen Y^*), thus activating the control system that leads, when possible, to the latter's neutralization. When the recognition fails the organism undergoes aggression from the antigens and often dies. When there is no recognition capacity the so-called autoimmune diseases can occur.

The box at the bottom right in Fig. 3.3 contains a conditional clause that allows the archive of objects-models to be updated in the event the object-model set as the objective is not recognized. Regarding this specific aspect, see Sect. 3.9.4.

3.3   Steering and Halt Control Systems

All the control systems we have examined so far have two important characteristics: they were constructed to achieve a fixed objective, Y^*, over the entire duration of the control—except in the case of occasional changes introduced by the governance—and the values of Y_t could be greater than or lesser than the objective. We can now relax these two conditions. Let us begin by considering the second condition.

Control systems—represented by the model in Fig. 3.1—in which Y can take on values that are above (e.g., water too hot) or below (water too cold) the objective Y^* (pleasant temperature) can be called "steering" control systems, or even "two-way" systems; the system acts to achieve the objective through positive or negative adjustments that converge toward Y^*, independently of the initial value of Y. Control systems for sound volume, temperature, direction, and speed are typically "two-way" control systems. The models developed for "two-way" systems can also

be applied to "one-way" systems, or "halt" systems, where the values of Y must tend toward but not exceed Y^*, in which case the control process would fail. The error can never change sign.

Let us consider, for example, having to pour a liquid, X_t, into a container (a glass, artificial basin) until level Y_t reaches the value objective Y^*. It is clear that for values $Y_t < Y^*$ the control system continues to increase the level until $Y_t = Y^*$; if, however, we do not stop the flow when the objective is achieved but allow $Y_t > Y^*$, then the system has not attained the desired control (the glass overflows and the basin floods the surrounding area).

"One-way" control is very common in all control contexts and is found in all situations where Y^* is by nature a non-exceedable limit or constraint. If by supplying light and heat to a greenhouse we can control the ripening state of strawberries before picking them, then if we exceed the objective we lose the harvest. If we want to reduce the content of a tank by removing the drain-plug, we cannot exceed the lower limit that is set as an objective. If we want to catch the bus that leaves at a certain hour and the control fails, the failure is signaled by the fact we have to wait for the next bus.

"Halt" control systems can also be represented with the same models as the others; the uniqueness of such systems merely requires a "precise regulation" of the chain of control of Y so that it does not exceed Y^* before we stop the control process. This implies the need to *measure* the error at the end of short intervals, and without delays, and to carry out the *regulation* without delays as well, thereby reducing the action rate "g" when the error falls below the critical level, thus moving from a normal to a "fine tuning" control. In fact, we all reduce the speed of our car to a minimum when we must park it near the wall in our garage in order to avoid dents, just as we reduce the flow of champagne (by adjusting the inclination of the bottle) when we reach the capacity level of the glass, or lower the flame when the roast appears to be done. Even my Mac possesses a "fine tuning" utility for brightness, called "Shades," which, contrary to the classical system of control, allows me to proceed in smaller increments to have a better control on the level of brightness.

In order not to risk exceeding the objective or limit, control systems, especially "Halt" control systems, may call for particular apparatuses that send the manager a particular "warning" of "nearness to the objective," above all when it is difficult to detect the error. We are all familiar with the beep-beep sound that warns us when we are near an obstacle while we are backing up, signaling a need for "fine tuning" our speed.

These apparatuses can be considered part of the broad class of *algedonic alerts* (from the Greek αλγος, pain and ηδος, pleasure), which associates a signal of *urgency* to the error $E(Y)$, even when this is not directly detected; the signal is usually accompanied by a sensation of pleasure or pain, so that the manager is moved to quickly eliminate it by controlling Y_t. If we touch an object which is too hot, a signal of pain makes us immediately withdraw our hand from the object, even if a thermometer to detect the temperature would have stopped us from touching the object in the first place.

The term *algedonic signal* has been used by Stafford Beer to indicate a form of control exercised by the environment on our behavior and objectives, through a system of reward or punishment. Each reward or punishment spurs us to modify our former objectives, and thus our behavior, which subsequently will again be judged by the environment to determine whether it merits new rewards or punishments (Widrow et al. 1973). Thus, algedonic alerts are particularly useful for understanding that the outside environment has detected errors, $E(Y)$, in our behavior and that we must urgently change our objectives and perhaps even the entire control system.

> Algedonic Regulation: literally, regulation by pain and pleasure, more generally, by rewards and punishments for the products rather than the behavior leading to it. E.g., people may be trained to perform a task by explaining to them the role their task plays within the larger system of which they are a part, but they may also be trained algedonically by a series of rewards and punishments that offer no such explanations. The algedonic regulator must have an image of the expected system of behavior but it restricts fluctuations not in the behavior of its parts but in their outputs (Krippendorff 1986).

> There is a last feature of the VSM to incorporate: it is extremely important, but there is not much to be said—so don't flick the page too soon. We were speaking of somnolence. It is an occupational hazard of System Five. After all, all those filters on the main axis … maybe Five will hear the whole organism droning on, and simply 'fall asleep'. For this reason, a special signal (I call it algedonic, for pain and pleasure) is always identifiable in viable systems. It divides the ascending signal—which we know to be entering the metasystemic filtration arrangements—coming from System One, and uses its own algedonic filter to decide whether or not *to alert system five*. The cry is 'wake up—danger!' (Beer 1995, p. 126).

3.4 Fixed- and Variable-Objective Systems (or Systems of "Pursuit")

Turning to the other condition regarding the type of variability of the objective (Sect. 3.3 above), we have already seen that we can also achieve objectives that vary over time in an intermittent and occasional manner, so as to form a path (Y_t^*) which the variable Y_t must follow, as shown for example in Figs. 2.4 and 2.8. I propose the following generalization: control systems can also be structured to achieve objectives that *systematically vary* over time: that is, *dynamic objectives*.

If the dynamics of the system occur at discrete instants, then the variation in the objective always precedes the variation in X_t and Y_t, so that, in principle, the system might not be able to achieve its objective. However, we can distinguish two cases:

1. The objective varies independently of the variables X_t and Y_t and has the following form: $Y_{t+1}^* = H(Y_t^*)$, where H has to be specified.

Under these assumptions, and agreeing that, for simplicity's sake, $H(Y_t^*) = Y_t^* + H_t$, the control system can be defined as *evolving*, since it tends to achieve an objective that varies, or evolves, for *exogenous reasons*; for example, in the control of the level of liquid in a container, which changes with each
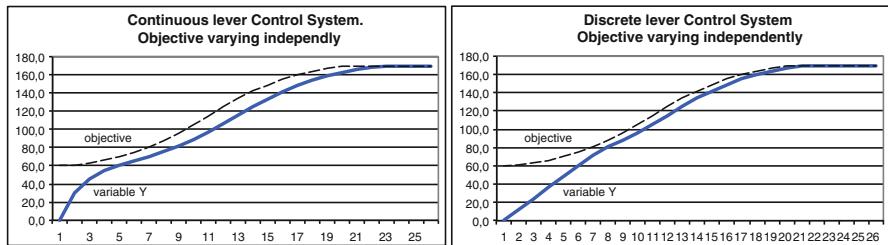
Fig. 3.4 Dynamics of a single-lever system of pursuit with an independently varying objective

fill up, or in many games where the players modify the position of an object-objective (a ball that changes direction when hit by other players).

The system will achieve its objective if the dynamics of the latter stabilize around a constant value for some period of time. Evolving systems are typical in living and social systems, in individuals and organizations, since they characterize the development phase. Figure 3.4 shows the dynamics of a single-lever system with a variable objective having the form:

$$Y_{t+1}^* = Y_t^* + H_t,$$

with "H" taking the values of $H = 0$ to $H = 10$, for $t = 0$ until $t = 10$, and the values from $H = 10$ to $H = 0$ for $t = 11$ until $t = 21$.

Since at $t = 0$, $Y_0^* = 60$, and since for $t > 21$ the objective has a constant value of 170, the system succeeds in carrying out the control with an assumed reaction time of $r(Y) = 1$; the longer the reaction time is, the more the variable representing the objective moves away from Y.

2. The objective varies as a function of $E(Y)_t$ (or X_t) according to the following form: $Y_{t+1}^* = G(Y_t^*, X_t, E(Y)_t)$, with G representing the "speed of flight" of Y_t^* (G can also be negative), which needs to be specified.

Under these assumptions, and also assuming, for simplicity's sake, that

$$Y_{t+1}^* = Y_t^* + GE(Y)_t$$

such systems can be defined as *systems of pursuit,* since it seems that Y_t tries to pursue the objective by reducing the error $E(Y)_t$, and that Y_t^* "observes" the controlled variable and adjusts its values to those of this variable, trying to "flee" from the pursuer by further increasing the distance by $GE(Y)_t$. On the one hand, the size of the variable Y_t depends on the overall amount of $E(Y)_t$, and on the other the size of the objective is, in turn, modified based on the coefficient G applied to $E(Y)_t$. The conditions that enable the pursuer, Y_t, to reach the pursued, Y_t^*, can easily be calculated, though they differ according to the type of *control lever*:

1. If the lever X_t varies *continuously* according to the fixed *action rate* "g," and the system operates with a fixed *reaction time* "r," then Y_t reaches Y_t^* if $G < 1/r$; in

fact, recalling (2.6) and (2.9) from Sect. 2.3 (omitting the time reference),
Y increases by $\Delta Y = Xg/r = Ehg/r = Eg$, while Y^* by $\Delta Y^* = EG$; if we want
$\Delta Y > \Delta Y^*$, then we must have $E/r > EG$; it follows that if $G < 1/r$, then the
pursuer will reach the pursued; if $G > 1/r$, the pursued flees and is never
reached; if $G = 1/r$ there is a position of deadlock.

2. If the lever X_t varies *discretely*, with a *fixed rate* "g," and the system has a
   reaction time $r = 1$, then Y_t reaches Y_t^* if $G < g/E_0$; the demonstration of this is
   analogous to the previous point.

Figure 3.5 shows various tests. In both cases we can modify both "g" and "G" so
as to produce dynamics of pursuit with a variable "distance," as shown in Test
3, where we assume that, upon seeing Y_t approach, Y^* increases his speed of flight
from $G = 33\ \%$ to $G = 60\ \%$.

The cases I have presented, as abstract as they may seem, are significant
regarding the situations of escalation we frequently encounter in real life: from
the pursuits and flights of cars or Olympic races to the pursuits and flights involving
predators and their prey in the struggle for survival (Sect. 7.4).

In fact, observing such processes, it is natural to think they are regulated by
systems of pursuit. The lion pursues the gazelle, but if this has a constant competitive advantage the lion will never catch up to it, as the following African proverb
clearly attests to:

> Every morning in Africa, a gazelle wakes up. It knows it must outrun the fastest lion or it
> will be killed. Every morning in Africa, a lion wakes up. It knows it must run faster than the
> slowest gazelle, or it will starve. It doesn't matter whether you're the lion or a gazelle: when
> the sun comes up, you'd better be running.

Here is a significant variant of the above:

> A lion wakes up each morning thinking, "All I've got to do today is run faster than the
> slowest antelope". An antelope wakes up thinking, "All I've got to do today is run faster
> than the fastest lion". A human wakes up thinking, "To hell with who's fastest, I'll outlast
> the bastards".

Systems of pursuit can give rise to various forms of *escalation*, both in a
biological context—as occurs, for example, in the evolutionary escalation of the
systems of attack and defense of predators and prey, respectively, as Darwin so well
described—and in a sociopolitical context. We can observe these tendencies, for
example, in arms escalations, feuds, wars, and violence, as well as in the economic
field, when firms develop their strategies to improve quality and productivity to win
the "war" of competition.

A special *system of pursuit* is the one that leads to continual progress in all the
scientific, technical, economic, and business sectors. A sports record (Y^*) is set and
immediately thousands of athletes train and compete to break it (objective); the
average level of preparation (variable X) and performance (variable Y) increases, so
that, sooner or later, someone sets a new record, thereby motivating thousands of
athletes to train and compete in the continual pursuit of an objective that always
seems to be just out of reach.

Test 1 Continuous lever $Y^*_0 = 60$ units $Y_0 = 0$ $E_0 = 60$ $g = 20; r = 2$ $G = 33\%$ **$1/r > G$**	
Test 2 Continuous lever $Y^*_0 = 60$ units $Y_0 = 0$ $E_0 = 60$ $g = 20; r = 2$ $G = 60\%$ **$1/r < G$**	
Test 3 Continuous lever $Y^*_0 = 60$ units $Y_0 = 0$ $g = 20; \ r = 2$ $G = 33\%, \ 0 \le n = 9$ $G = 60\%, \ n > 9$ **$1/r > G, \ 0 \le n = 9$** **$1/r < G, \ n < 9$**	
Test 4 Discrete lever $Y^*_0 = 60$ units $Y_0 = 0$ $E_0 = 60$ $g = 30; \ r = 1$ $G = 45\%$ **$g/E_0 > G$**	
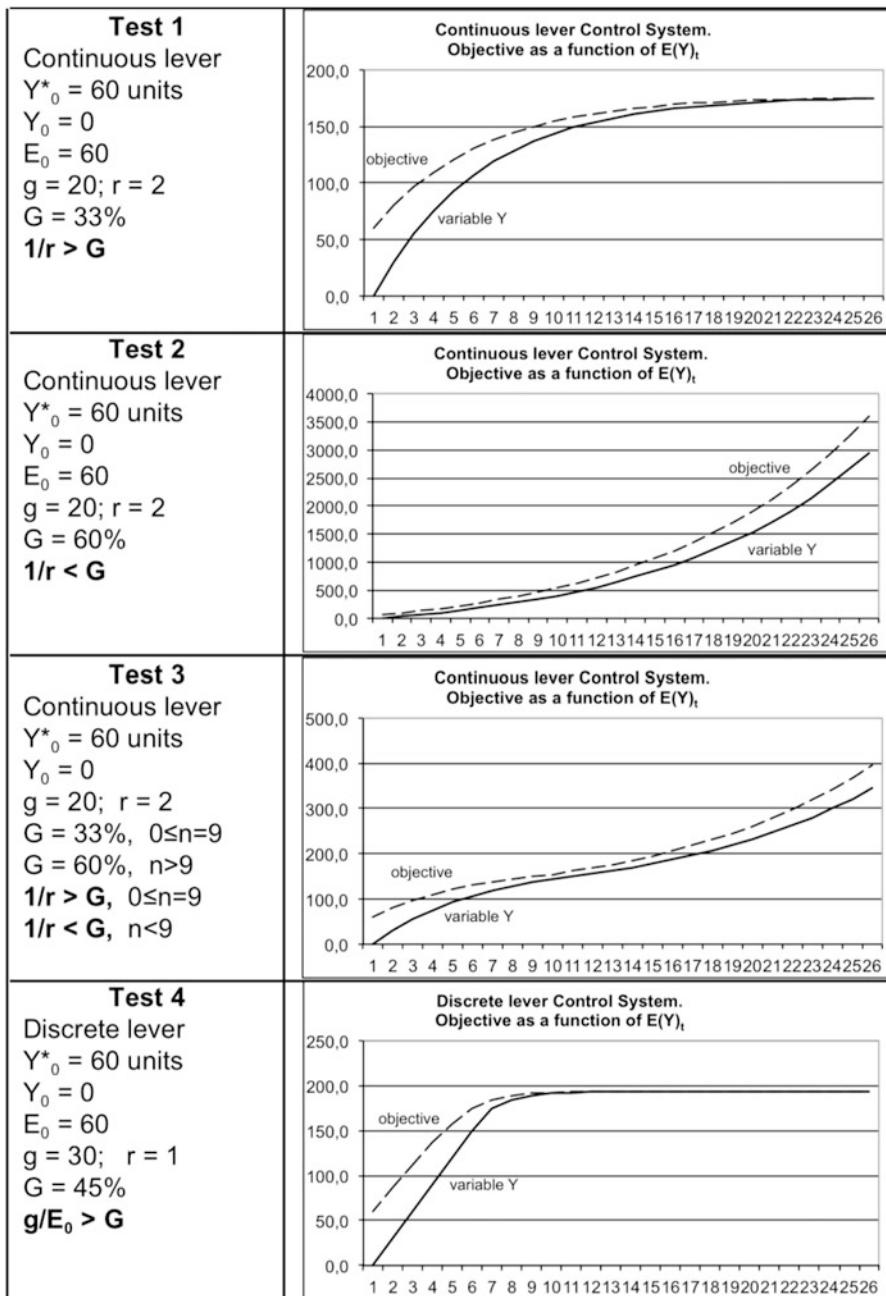
Fig. 3.5 Dynamics of a single-lever system of pursuit with an objective varying as a function of $E(Y)_t$

A similar process occurs when a new scientific discovery (Y^*) induces thousands of researchers to broaden their research (variable X), which soon gives rise to new discoveries (variable Y). The level of scientific knowledge (objective) continually progresses. How else can we explain the dynamics of progress and well-being if not as the result of the action of control systems with variable objectives? I shall return to a discussion of the above-mentioned aspects in Sect. 5.7.

3.5 Collision, Anti-collision, and Alignment Systems

Systems of pursuit can be seen in systems of *collision, anti-collision*, and *alignment*, whose logic can be summed up as follows. Assume there are two Control Systems, A and B, which interfere with each other, in the sense that one of the two, say B, has the objective of achieving and maintaining a certain velocity, while A's objective is to achieve and maintain B's velocity, increased or reduced by a constant P (in the simplest-case scenario). If $P > 0$, then A will catch up to and exceed B's speed. If A and B are in single file (same route objective) a collision will occur. If A is designed to collide with B then we have a *collision system*. If A is following B, then A is a system of pursuit with respect to B. If A precedes B, then A is a system of flight from B. Clear examples of such systems are predators that must pursue their prey and catch up to it, that is, collide with the prey, or prey that must flee from their predators by increasing their speed. If $P = 0$ then A will maintain the same speed as B; if they are vertically aligned we get the *alignment* of A and B, where A follows B. If $P < 0$ then A will maintain its speed below that of B by the constant P, which has a negative value. System A is an anti-collision system, since A is lined up behind B but, in order to avoid colliding with B, A reduces its speed with respect to B.

We are used to considering collision or anti-collision processes as spatial processes carried out by systems whose objectives are defined as dynamics in space. There is no contradiction here with the model just described, since, if we assume that the systems A and B are vertically aligned—that is, they are on a trajectory along a curve, however this is defined—then it is easy to translate their differences in speed (as well as the constant P) in terms of distances along the spatial trajectory the systems themselves are moving along. It would also be possible to carry out a control of the spatial trajectory as well as the velocities in order to guarantee a collision or to avoid one. Nevertheless, we would have to consider two objectives contemporaneously (spatial position and speed) and several levers. We shall discuss these aspects in Sect. 4.6.

3.6 Tendential and Combinatory Control Systems

I shall conclude this *short typology* by mentioning two special classes of control system that will be useful when we examine combinatory systems in some detail (Sect. 7.8).

The first class are systems that can be called *tendential* control systems. In many feedforward control systems, the control variable X acts *randomly*, in the sense that, although it is calculated in order not to produce a gap between Y and the objective, Y^* (which, for simplicity's sake, we shall consider constant over time), it produces values of Y_t that do not permit the objective to be pursued. This is due to the work of unknown and uncontrollable *disturbance* factors that intervene, which we can define as *causal* in all respects. If the feedforward control system operates over a certain number of cycles—for example, an archery champion trying to hit the bull's-eye—we can assume that the system manager (the athlete) gains more and more experience and determines the values of X in an increasingly more accurate fashion, thereby increasing as well the precision of Y with respect to the Y^*.

To evaluate the amount of control of such a system—that is, to understand if over time the system will move closer to the objective—we must evaluate not the gaps between each single Y_t and Y^*, but the *trend* calculated using a set of values, Y_t, based on a static procedure we have chosen to describe the dynamics, since Y_t can be treated as a *time series* in all respects. I will only briefly mention the following possibilities, suggesting the reader consult a statistics text for the applied techniques:

(a) Evaluate the *trend* by simply calculating the moving arithmetic average for all values of Y_t from $t = 0$ to the interval of the last observation; the closer the average is to the objective, the more we can assume that the system is "tendentially" reaching the objective.

(b) If we need to make a progressive evaluation, then rather than using the average for all values of Y, it could be useful to estimate the trend by calculating the *moving average* using an appropriate interval of values centered on Y_t; since there are many types of moving averages—*simple moving average*, *weighted moving average*, *exponential moving average,* to name but a few—we must choose the type of moving average that is most significant for expressing the *speed of adjustment* of Y to the objective, taking into account the nature of the variables and the length of the various control cycles of the system;

(c) Evaluate the *trend* using least-squares techniques; in the simplest case, the trend may be expressed by estimating the line which minimizes the square of the variance of each Y_t with respect to Y^* for the interval of observation and repositions the slope of the trend line through the given data points of Y_t.

These statistical techniques to detect the dynamics of the error and its movement toward zero are widely used in performance quality control for sports, organizations, and companies.

Figure 3.6 represents the control chart of a system of regulation that, acting on a *control lever* X, seeks to maintain the dynamics of the values of the controlled variable Y_t, which expresses the quality of a generic production process, within the objective values $Y^* = 130 \pm 5$. The control chart indicates the moving averages of 5 and 9 values to test whether or not there is movement toward the objective. We see from an initial visual analysis that from the period $t = 16$ onward the values of the variable Y_t tend to stabilize around values close to the objective; this shows that the control actions on the technical variables X were successful.
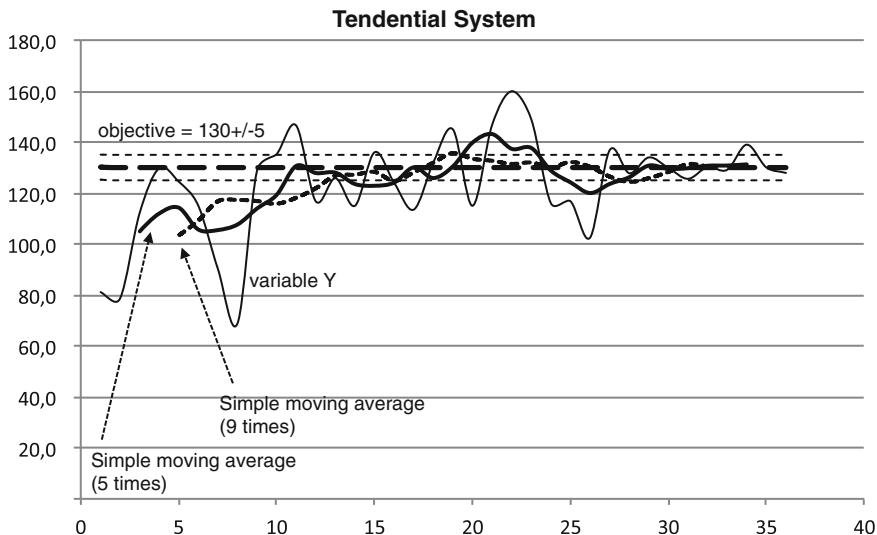
Fig. 3.6 Dynamics of an abstract tendential system

The second type of system I shall refer to as *combinatory* control systems. Their unique characteristic comes from the fact that the controlled variable Y is the composite result of the combined action (the meaning of which will be defined below) of a group of individuals (animals, persons, or other) whose collective action, X, produces the values of Y that permit the objective, Y^*, to be achieved.

The gap between Y and Y^* results in a change in the behavior of the individuals—no matter whether uniform or varied, contemporaneous, synchronized, or at different time periods—so that the "combination of the behaviors," X, produces a new value for Y.

Combinatory systems usually also act as *tendential systems*, in that in the combined action of various individuals the values of the variable Y must be tendentially evaluated in their movement toward the objective.

To understand the logic of combinatory systems it is useful to refer to Fig. 3.7, which shows the values $[y_t^1, y_t^2, \ldots, y_t^N]$ lined up in rows, obtained from $N = 10$ agents over $T = 20$ cycles as a function of a control lever X (not shown in the table). The system's response in terms of Y_t is determined as the average of the individual values of the $N = 10$ agents and is presented in the "average" column. Even without a detailed examination, a simple reading of the data indicates that, starting from period $t = 10$, a specific control has been carried out on X (not shown in the table) that gradually eliminated the variance. Naturally, since we are dealing with a *tendential system* the variance in this specific example is never completely eliminated; although it returns in subsequent periods, it nevertheless remains at moderate values. Combinatory control systems are quite common, and they will be dealt with more at length in Sect. 7.8.

cycles	A(1)	A(2)	A(3)	A(4)	A(5)	A(6)	A(7)	A(8)	A(9)	A(10)	mean = Y	objective = Y*	Error = Y* - Y	
t=1	153	155	151	150	155	150	155	149	148	151	152	160	- 8	Action on X
t=2	154	155	149	155	154	152	156	149	149	153	153	160	- 7	
t=3	154	159	149	152	158	152	156	147	149	153	153	160	- 7	
t=4	156	159	152	156	158	152	156	150	151	154	154	160	- 6	
t=5	156	160	153	156	158	152	156	153	151	154	155	160	- 5	
t=6	156	160	153	156	158	155	156	155	151	158	156	160	- 4	
t=7	157	160	155	156	162	157	156	155	151	158	157	160	- 3	
t=8	159	163	155	157	162	157	157	157	151	158	158	160	- 2	
t=9	159	165	157	159	162	157	158	157	155	158	159	160	- 1	
t=10	163	165	160	159	162	160	158	157	155	158	160	160	- 0	maintaining control
t=11	165	168	160	159	162	160	161	158	156	159	161	160	1	
t=12	162	160	164	161	162	160	163	158	156	163	161	160	1	
t=13	160	160	162	161	162	160	160	160	162	163	161	160	1	
t=14	158	158	162	161	162	162	158	158	162	163	160	160	0	
t=15	157	165	157	158	167	159	157	158	162	166	161	160	1	
t=16	157	161	157	160	159	159	159	160	161	166	160	160	0	
t=17	155	163	160	158	159	158	157	155	157	166	159	160	- 1	
t=18	159	162	161	160	161	160	159	155	154	166	160	160	- 0	
t=19	161	162	162	160	159	162	159	155	154	166	160	160	-	
t=20	162	164	157	157	157	163	158	159	156	160	159	160	- 1	

Fig. 3.7 Logic of combinatory control systems

3.7 Parallel or Serial Connections

Until now we have considered the control of the variable Y_t in its movement toward the objective Y^* as a process carried out by a single control system. We can immediately extend this logic to cases in which the control of Y_t is carried out by a plurality of control systems connected *in parallel* or *serially*.

The *parallel control* occurs when changes in Y_t require the action of a lever, X_t, which cannot be activated by a single control system. Thus, *partial* control systems are set in motion which are small in scale and capable of producing an overall effect equal to that of a larger-sized control system.

For simplicity' sake, let us assume that only *two* partial systems, A1 and A2, are activated; then, the *parallel* connection can be achieved in two ways.

(a) *Decomposition*: the objective Y^{A*} is decomposed (that is, divided) into two subobjectives, Y^{A1*} and Y^{A2*} (which are not necessarily equal), so that $Y^{A1*} + Y^{A2*} = i^{A*}$; these partial objectives are assigned to two *parallel partial systems*, A1 and A2.
(b) *Composition* of the lever X^A that controls Y^A by summing the values of the levers, X^{A1} and X^{A2}, of the two subsystems.

Form (a) of parallel control is purely additive, since we sum the effects produced by the partial objectives. Figure 3.8 shows a possible simple parallel control system based on the decomposition approach; in this case the system A must control the variable Y^A through the *partial* systems A1 and A2, which are connected systems (Sect. 2.14).

The objective of A, $Y^{A*} = 350$, is divided into two subobjectives, $Y^{A1*} = 210$ and $Y^{A2*} = 140$ (other rules for subdivision would be equally admissible). The system A2 is "slower" than A1 since it is characterized by a reaction time of $R = 3$ and an action (power) rate equal to 1.
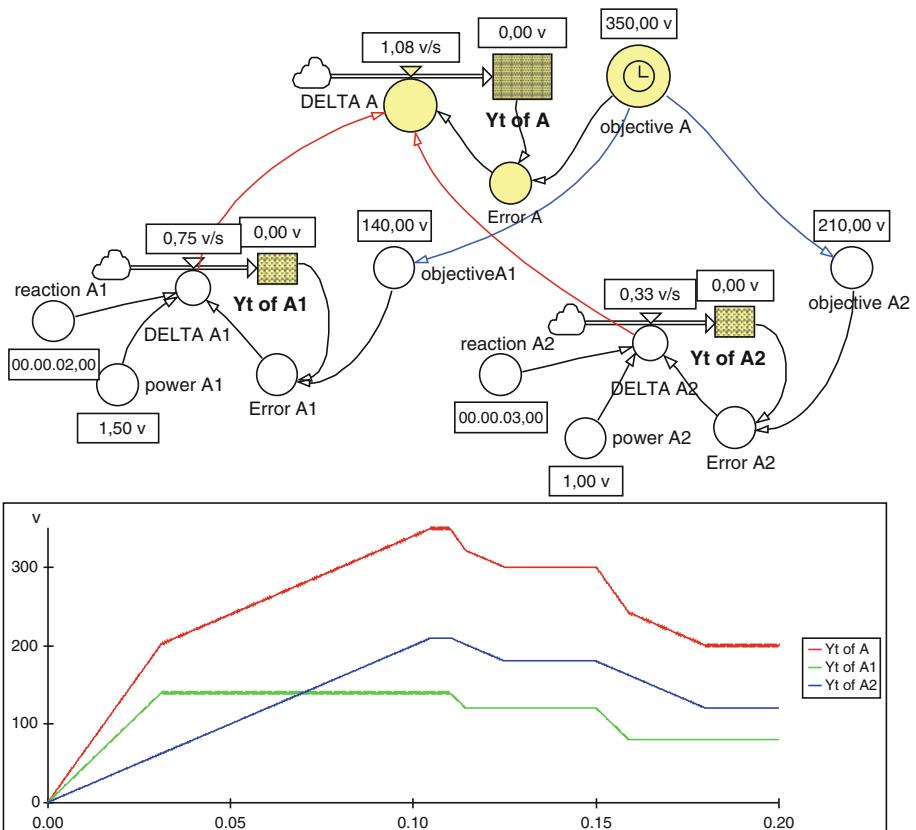
Fig. 3.8 Parallel systems based on the decomposition of $Y*$

Parallel control through the *decomposition* of the objective is rather common in both the socioeconomic and corporate environments. A firm could divide up the general revenue objective into partial objectives to assign to the various geographic divisions. The government, whose objective is to guarantee its citizens a certain quantity of health and educational services, should necessarily assign parts of these objectives to the various health and educational units in the country.

Control by linking partial parallel systems can also occur through the *composition* of the control lever X, as shown in the second case above. Unlike the first case, here the objective Y^A* is not broken up, since we assume it must be achieved only through system A, on the condition the latter is able to assign a sufficient value to its own lever, X_t^A. When it is not possible or convenient to directly give a value to X_t^A of system A, then two partial systems, A1 and A2, could be activated in order to independently produce the partial values of the lever X_t^{A1} and X_t^{A2}, so that their sum (composition) result in: $X_t^A = X_t^{A1} + X_t^{A2}$.
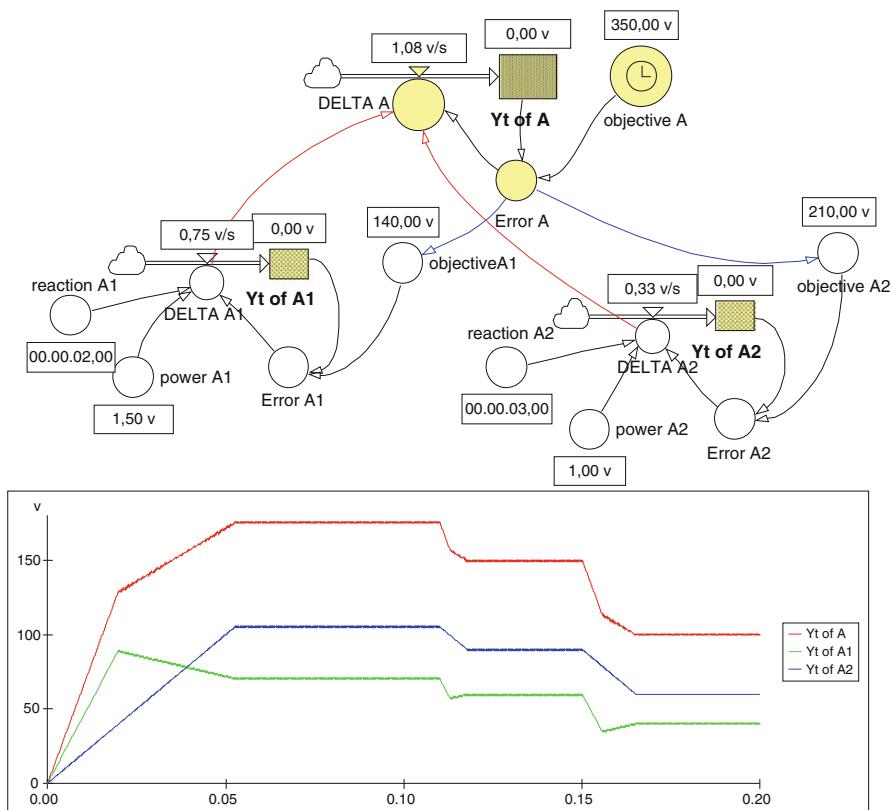
Fig. 3.9 Parallel systems from the composition of X

This form of segmentation is also quite common. If a force of 100 men is needed to move an obelisk, each of these men represents a partial control system that does not move the obelisk by even a hundredth part, but simply produces one-hundredth of the force necessary to place the obelisk in the correct position.

However, with parallel control systems through *composition* of the X_t^A we need to clarify what objectives $Y^{A1}*$ and $Y^{A2}*$ to assign to A1 and A1.

A simple solution is to assign to A1 and A2 an objective corresponding to a share, E^{A1} and E^{A2}, of the error of the main system, $E^A(Y) = Y^A* - Y_t^A$. As a result, in trying to produce values $Y^{A1}*$ and $Y^{A2}*$ to achieve E^{A1} and E^{A2}, A1 and A2 in fact end up producing a value of the A lever, $X_t^A = X^{A1*} + X^{A2*}$, which is necessary to achieve the general objective Y^A*, thereby eliminating the overall system error $E^A(Y)$, as can be seen in Fig. 3.9. In this figure, even with the same parameters of the model in Fig. 3.8, we note a different movement toward the achievement of Y^A*. In fact, the two systems, A1 and A2, in Fig. 3.9 achieve their objectives contemporaneously, since they uniformly eliminate over time the shares

E^{A1} and E^{A2} assigned to them; precisely for this reason the control ends when E^{A1} and E^{A2} become zero.

Let us now consider the structures that are formed when several control systems operate *serially*. Two systems, A (upstream) and B (downstream) are serially connected when the variable controlled by A, Y^A, becomes the lever X^B of the second system in order to achieve the objective Y^{B*}. Thus, the following causal chain is formed: $[X^A \rightarrow Y^A = X^B \rightarrow Y^B \rightarrow Y^{B*}]$, which clearly shows that Y^B is in fact controlled by the lever X^A of the system upstream. The objective to assign to system A can be represented by the values that must be produced in B's lever: that is, $Y^{A*} = X^{B*}$.

If it is possible to calculate in advance the value X^{B*} needed to achieve the objective of B, then Y^{A*} is a *fixed* objective; otherwise it is *variable*, and A continues to operate until the system *downstream* has not eliminated its error. In both cases the objective of the system *upstream* is assigned by the control apparatus of the system *downstream*.

Serial control is the rule, not the exception. If we apply the FIRST RULE of systems thinking (Sect. 1.2) and zoom in, we can easily see that most control systems can be *serially* coupled to another instrumental system whose function is to produce the variations in the first system's lever.

If we instead zoom out we can easily realize that each control system can be an element, in *parallel* or *serially*, for a larger system. In order to drive a car we have to turn the front wheel (downstream system) appropriately; however, this depends on the rotation of the steering wheel, which in turn depends on the "torque" produced by our arm (upstream system). When we, along with our colleagues, achieve the objectives assigned to us at work, the department also achieves its objectives; if all the departments achieve their objectives then the organization as a whole does so as well.

When we observe control systems from a systems thinking perspective and try to identify the relations between X and Y, having in mind only a single objective $Y*$, we are at a summary level; that is, we do not zoom in and explicitly consider whether this system derives from the composition of parallel partial control systems or whether other variables operate between X and Y, each of which is controlled by serially arranged partial systems. In order to treat the topic in a manageable way, in this *section* we have only considered two partial subsystems, A1 and A2 or B. More generally, we have to keep in mind that both *parallel* and *serial* arrangements can be formed by any number of partial systems.

Finally, we must consider that *parallel* and *serial* structures can be joined together to form complex control structures that characterize biological, social, and organizational contexts. These aspects will be taken up in Part II of the book.

3.8 Holarchies of Control Systems

The points developed in the previous *section* can be generalized from a holonic perspective (Mella 2009). When we observe control systems from a systems thinking point of view and try to identify the relations between X and Y, having a single

objective, Y^*, in mind, we are treating the topic in a concise manner; the control system is viewed as an autonomous entity. Systems thinking teaches us, however, that we must get accustomed to thinking of every control system as interacting with other systems at the same level, or at a superordinate or subordinate one. In addition to analyzing the *serial* and *parallel* connections, which produce a mono-directional interaction among control systems, we can and must extend our observation and consider control systems as classes of observed entities arranged in *hierarchical levels*. This necessity is made clear by reflecting on the fact that many control systems, though apparently acting independently, are nevertheless necessary for the functioning of others at both a superior and inferior level; this is true for any level of observation. In other words, if we zoom in and out at the same time it appears that everywhere there is a holonic hierarchy, a *holarchy*, among control systems.

We become aware of the various levels of control systems as soon as we realize we can control the actions of our body (move ourselves, feed ourselves, etc.) because our organs function correctly thanks to their control systems; these organs, in turn, function without errors thanks to the control systems in the cells that make up the tissue in our bodies; the cells function as a result of the control systems which regulate their autopoiesis.

It is clear that I can control my fingers on a keyboard to write this sentence, since I control my arm, shoulder, and entire muscular-skeletal apparatus; moreover, the flow of words can pour out from the keyboard because the recognition system, thanks to the neural system, identifies the appropriate words from the catalogue of my knowledge (grammar and vocabulary), arranging them in the correct sequence. However, this implies that the areas of the brain that concern memory and various other functions are being controlled; and, at a lower level, the individual neurons are controlled, which in turn control the impulses that arrive from "upstream" to discharge other impulses "downstream." It is equally clear that the control of the direction and speed of our car is possible since its various organs and components—amount of gas in the tank, turbochargers, electronic system, tires, etc.—are linked to specific control systems.

Even the control system that seeks to keep pollution under certain thresholds around the globe implies emissions control in large urban areas, and, at increasingly lower levels, in cities and energy-providing companies, which limit individual emissions by controlling the energy consumption of plants and the maintenance state of facilities. How can we control infant mortality rates if we do not activate controls on research and health structures (research labs, obstetric and neonatology departments) and on families (hygiene and diet)? How can we control health and family-related structures if we do not control the level of preparation and education of their personnel? In short, while taking nothing away from the circular interconnections among control systems, it is clear that there are various levels of control systems that constitute a *holarchy*.

The notion of *holarchy* comes from the concept of *holon*, coined by Arthur Koestler (1967, 1978), who viewed the *holon* as a *Janus-faced entity*: if it observes its own *interior* it considers itself a *whole* formed by (containing) subordinate *parts*; if it observes its *exterior* it considers itself a *part* or *element* of (contained in) a

vaster *whole*. In any event, it sees itself as a *self-reliant* and *unique* entity that tries both to survive (it is a *viable system*) and to integrate with other holons:

> These sub-wholes—or 'holons', as I have proposed to call them—are Janus-faced entities which display both the independent properties of wholes and the dependent properties of parts. Each holon must preserve and assert its autonomy, otherwise the organism would lose its articulation and dissolve into an amorphous mass—but at the same time the holon must remain subordinate to the demands of the (existing or evolving) whole (Koestler 1972, pp. 111–112).

Each holon includes those from lower levels, but it cannot be reduced to these; it transcends them at the same time that it includes them, and it has emerging properties (Edwards 2003).

Due to their Janus-faced nature holons must necessarily be connected to other holons in a typical *vertical arborizing structure* known as a *holarchy,* which can be viewed as a *multi-layer* system (multi-strata) (in the sense of Mesarovic et al. 1970) or *multi-level* system, with a *tree structure* (Pichler 2000). Each holon is a *head* holon for the subtended part of the branch and a *member* holon for the upper part. It is relevant to observe that holarchies are not holons but arrangements of holons that represent conceptual entities whose function is to bring out the essentiality of the vertical interactions among holons.

The completeness principle (Mella 2009) must, in any event, apply in the multi-layer holarchy: each subordinate level represents holons which are less extensive and recompressed into the holons at the superordinate level (arborization effect), with the understanding that all the base holons must be included in the final holon.

Koestler (1967, p. 344) defines *output holarchies*—or *descendent* holarchies— as those that operate according to the *trigger-release principle*. In such holarchies the *top holon* carries out its own processes thanks solely to the activities of the subordinate holons, which the top holon itself coordinates by sending information about the activities the subordinate ones must undertake. He defines *input holarchies*—or *ascendant* holarchies—as those holarchies that operate according to the *filtering principle*; they produce progressive syntheses from the subordinate to the superordinate levels, as if, at each level, the holons filtered and synthesized the inputs from the subordinate holons.

We can easily note that the holarchies of control systems are at the same time both input and output holarchies; *they are descendent relative to the objectives and ascendant relative to the variables under control*. The top-holon control system "launches" its objective, Y^*, as information to lower-level control systems which, in turn, "launch" subobjectives to even lower levels. On the other hand, control system holarchies are also *ascendant*, in the sense that the values for Y (or X) at a certain level are syntheses of lower levels that are filtered (synthesized) to higher levels. The close connection between control systems for the Y variables at different levels (ascendant connection) and the Y^* objectives (descendent connection) is a general feature of control system holarchies; this connection fully reflects the general principle, well outlined by Koestler, of the behavioral complexity of holons:

> Hierarchies can be regarded as 'vertically arborising structures whose branches interlock with those of other hierarchies at a multiplicity of levels and form 'horizontal' networks:

arborization and reticulation are complementary principles in the architecture of organisms and societies (Koestler 1967, p. 345).

This feature is shown also in the model in Fig. 3.10; while the basic control systems have irregular dynamics, their aggregation produces less regular dynamics as we proceed upward toward higher-level control systems. Figure 3.10 presents a simple control system holarchy formed by partial systems, arranged over three levels according to a typically holonic arrangement set out in a descending *series*; the main objective of the *top holon* (A), $Y^* = 360$ units, is broken up into several percentages (which I have introduced for the simulation); the percentage shares of the main objective are assigned to the subsystems, which are arranged in *parallel* at each level. The control system at level A represents an accumulator of the variable Y for the sublevels A1 and A2, which in turn accumulate the values of the variable Y from the 5 base holons (A11, A12, A21, A22, A23). Each variable Y of each system and each level acts to achieve its particular objective and, moving up through the holarchy, to allow the objective of the top holon (A) to be achieved. The *upper* part of Fig. 3.10 indicates the initial situation and the subdivision of the main objective; the *lower* part shows the results of the simulation after 60 iterations. Figure 3.11 presents Y's movement toward the objectives; we can see that, despite the fact the initial objective was subject to two exogenous variations, the system correctly controls the Y variables.

From a *bottom-up* perspective, the values of the Y variable at level A derive from the values produced at the lower levels; moreover, these values, from a top-down perspective, are conditioned by the holons at the superordinate levels, since level A indicates the objectives to achieve at the lower levels, which for simplicity's sake are set equal to a proportional share of the objective of the higher-level holon. In this twofold relation, descendent (Y^*) and ascendant (Y), we obtain a co-evolving dynamics among the control systems at the various levels, following the typical relation that Ken Wilber, in his *Twenty Tenets* (Meyerhoff 2005), clearly describes:

> 10. Holarchies co-evolve. The micro is always within the macro (all agency is agency in communion). 11. The micro is in relational exchange with macro at all levels of its depth (Wilber 1995, pp. 34–35).

Such conditionings not only act in an ascending direction but reveal their effects in a descending one as well. This means that in a control system holarchy each system at a given level acts together to condition the higher-level systems (following one of the arrangements in Figs. 3.8 and 3.9); however, at the same time it receives from the latter appropriate inputs that guide its own behavior. These inputs are then also transmitted in descending fashion along the lower branch, thanks to the coordinating capacity the system has in relation to those systems composing the subtended branch. Moreover, as Wilber explicitly notes, since every holon at a certain level (and this is particularly true for holonically ordered control systems) includes subordinate holons in its own structure, even while it transcends them, in order to survive it must preserve and regenerate these, since it depends on them for its existence. This makes it more likely that lower-level systems will be maintained and consolidated over time.
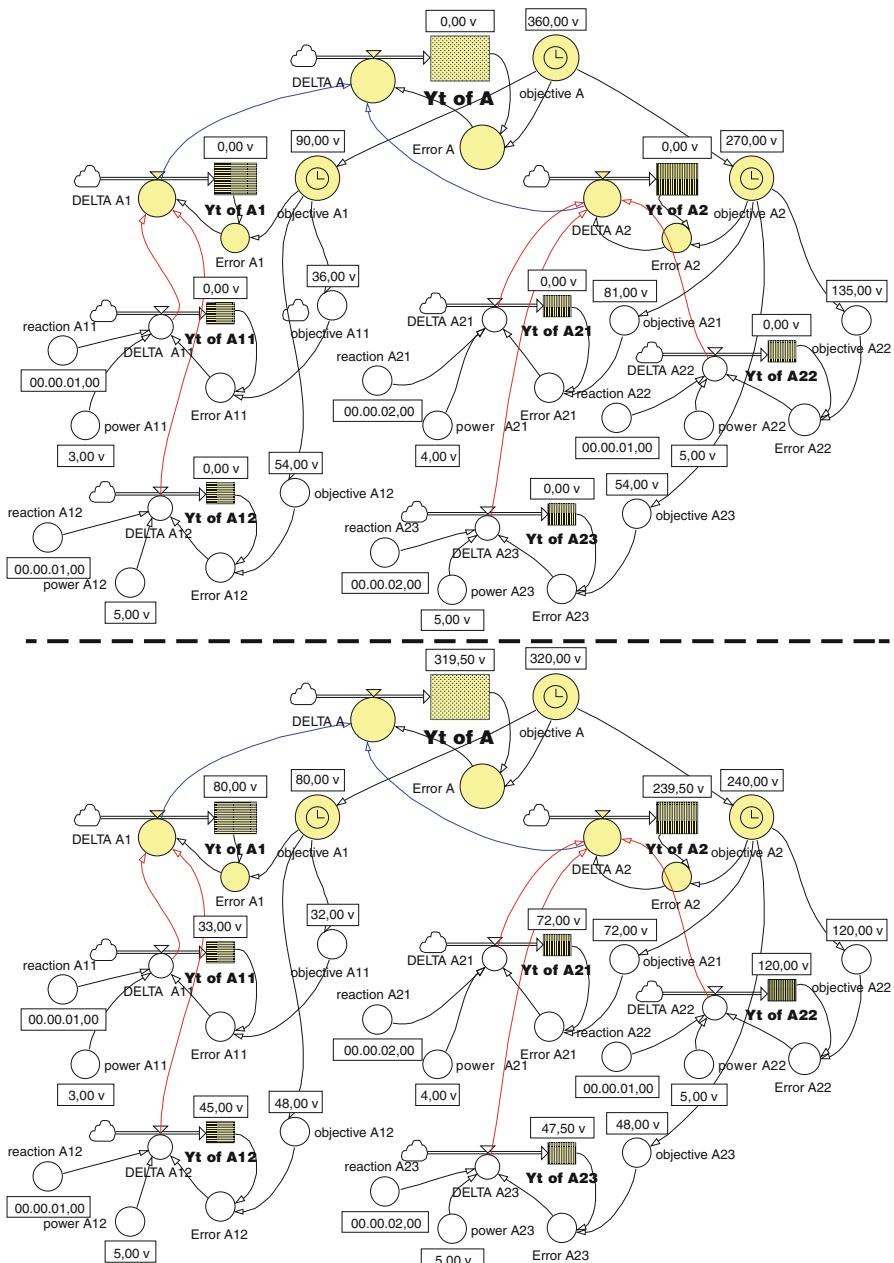
Fig. 3.10 Control system with three levels, based on a holonic arrangement
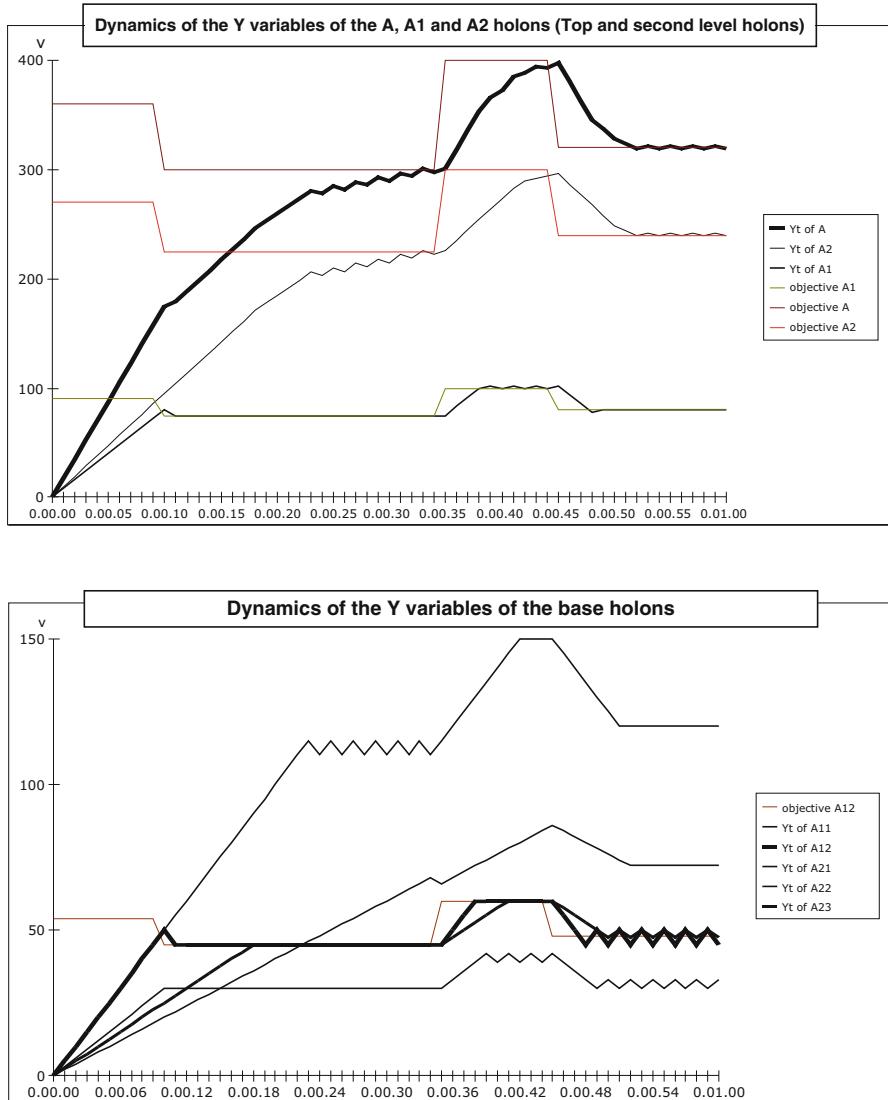
Fig. 3.11 Dynamics of the variables to control (Y) in a holonic system

From an *ascendant* observational perspective, by gradually zooming in we can verify that a miniscule pixel is functioning because a micro control system provides it with a certain state, since a superordinate control system makes this state necessary in order to adjust the state of the screen. However, the state of the screen derives in turn from the control objectives of the software operating at that moment, which in turn responds to the control needs of the person using it, who in turn is influenced by the control needs of the group in which he provides his activities,

which in turn are influenced by the control system represented by the operating center in which the group functions. This control system is part of a vaster control system of the department, the division and the entire organization, which in turn is under the control of other superordinate systems.

We can make similar considerations for all living organisms and for social organizations. For example, in living organisms the macro control of organs depends on the micro controls of the constituent cells; in a plant, each cell in the root produces a process controlled from a higher level, but the Y^{ij} variables of the control systems of cells are synthesized into the Y^j variables in the (increasingly larger) branching in the roots, eventually becoming the Y variable of the entire organ, represented by the roots as a whole. It is clear that the activities in the roots condition those in the other parts of the plant; however, it is equally clear that the operating objectives of the roots and branches, down to the individual cells, do not determine the higher-level objectives but, on the contrary, derive from the dietary and survival constraints of the plant as a whole.

Another example: the control of global warming implies emissions control by individual countries, which reverberates down to the control of the individual regions, companies or families, and the individual machines and their components that produce the substances feeding the *greenhouse effect*.

It surely is not the faster beating of the heart that spurs us to run in order to increase our blood pressure, blood circulation, and oxygenation; rather the opposite is true. Our objective to run imposes a control on the organs; the muscles need energy, there must be ample blood flow, the lungs must produce more oxygen, and thus the heart must beat faster, at levels determined by our physiological programming. Thus the heart is regulated by higher-level controls and, as a single unit, controls the contraction of its muscle fibers, which determine the movement of the cells of which it is composed. This may be a simplified view of things, but it is obvious that it is never the heart beating fast that forces us to run, but just the opposite: it is our running that forces our organism to also control our heartbeat. Only if the latter becomes anomalous and is not due to any special physiological need does the heart "spur" us to set in motion pharmacological or electronic control systems to bring our heartbeat "under control" according to our physiological objectives or constraints.

3.9 Complementary Material

3.9.1 Some Well-Known Cybernetic Systems

As observed in Sect. 3.1, cybernetics deals with automatic mechanical, biological, and social systems. In the collective mind cybernetic systems are usually thought of as mechanical systems that can control the variable Y by automatically regulating X, as we can clearly understand from the following quote:
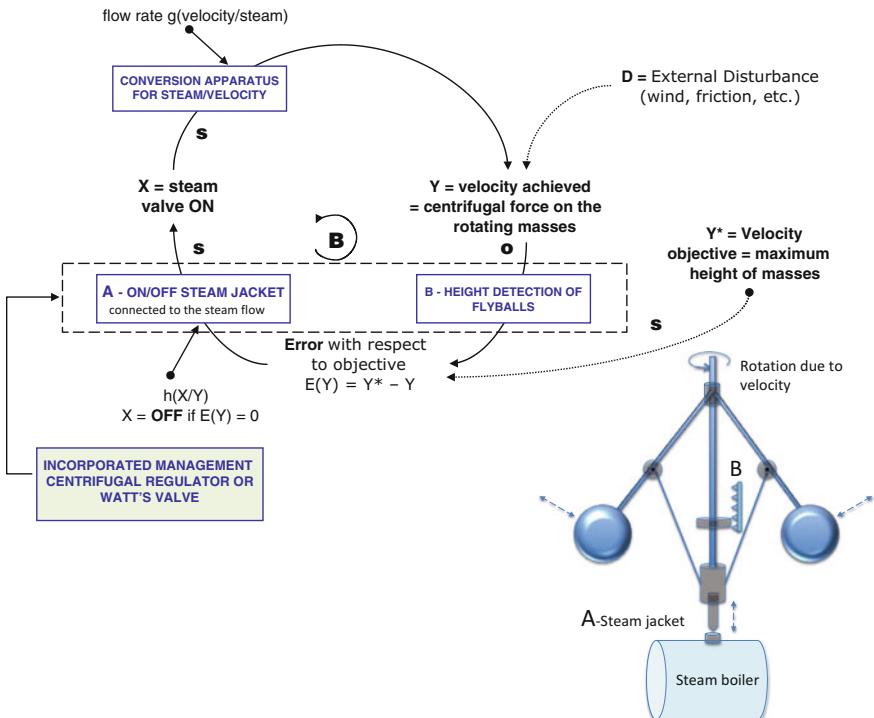
Fig. 3.12 Automatic control system. Watt's centrifugal regulator

Watt's governor was a superb example of feedback control. Feedback controllers, mechanisms that sense a discrepancy and correct it, are absolutely shot through our world today. We go through hardly an hour of any day without using feedback devices—float valves in our toilets, thermostats in our rooms, pressure-control valves and carburetion electronics in our cars.

[...] That sort of thing was common in the Hellenistic world. One of the first feedback devices was the water-clock flow regulator. The 3rd-century BC engineer Ktsebios made the ancient water-clock into an accurate timekeeper by inventing a float stopper to regulate a constant flow of water into the indicator tank (Lienhard 1994).

There has been rapid progress in mechanical cybernetic systems thanks to the ideas from engineers and technicians, which have succeeded in replacing the human manager with a mechanical one. The history of control systems and cybernetic systems has mainly been "written" by engineers who have designed and produced automatically controlled mechanical systems. Among the many examples presented below are those of models of the main automatic control apparatuses, indicated in the quote, which are now part of the history of cybernetic systems.

Figure 3.12 shows Watt's *Centrifugal Regulator* (otherwise known as the Centrifugal Speed Control Device, Flyball Governor, or Watt's valve) invented by Matthew Boulton and James Watt and produced in 1788; it served to automatically
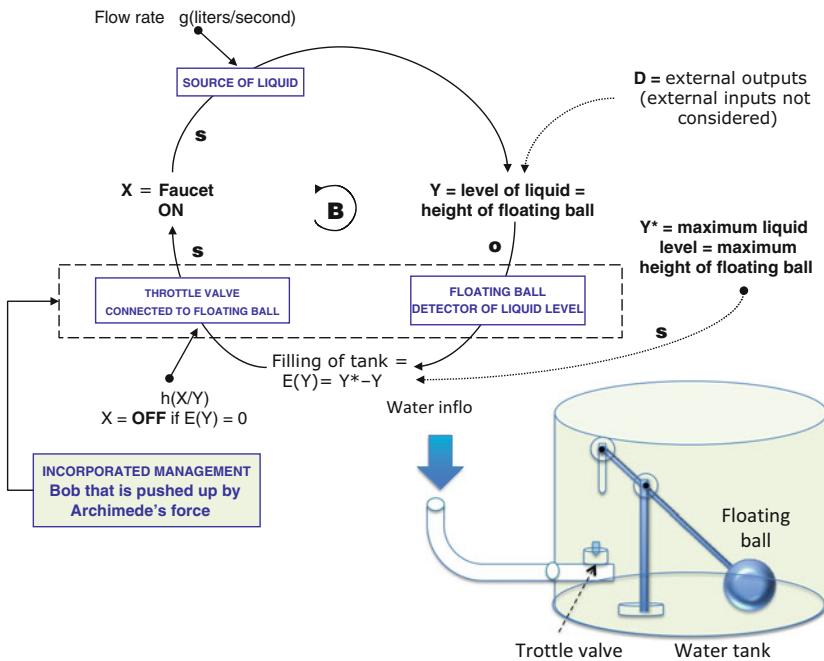
Fig. 3.13 Automatic control system. Floating ball regulator

control the speed of the steam locomotive and was applied to the control of the rotation of the driveshaft in all steam engines as well as windmills.

Watt's regulator is a valve that regulates the on/off state of the flow of steam that generates pressure in the piston. This is coupled to a mechanical speed detector made up of two oscillating spheres mounted on the spindle that rotates when activated by the driveshaft. When the speed of the motor increases, the oscillating and rotating spheres rise due to the action of the centrifugal force, thereby raising the axes linked to the on/off valve that closes off the steam when the speed raises the levers (due to centrifugal force) to the limit the governor has chosen as the speed objective. The lack of steam diminishes the pressure in the piston, which reduces the speed, thus lowering the spheres and once again opening the steam valve to increase the pressure and speed.

Figure 3.13 shows a system for the automatic regulation of the level of liquid in a container based on a principle entirely similar to that in Fig. 3.12. The level detector is represented by a bob that is pushed up by Archimede's force, while in Fig. 3.12 the two rotating balls were pushed up by a centrifugal force. A flow of liquid, constant or irregular, raises the level in the container, which in turn raises the bob. When the bob reaches the maximum desired level a butterfly valve is closed by the force of buoyancy, which stops the flow of liquid. The automatic system in Fig. 3.13 can be improved by also controlling the outflow of the liquid in the tank through a valve opener to eliminate the excess liquid in the event liquid flows in from an

outside disturbance. This system will be examined in the next chapter, as the addition of this second control lever leads to a two-lever system (Fig. 4.1).

Since these systems were conceived and created in previous centuries, it would seem they would be outdated. Yet their main principle is still valid; the automatic control is achieved through more modern detection and regulation systems, which are supported by electric engines commanded by relays. The *Flyball Governor* has been greatly simplified, since the operation of the *flyballs* has been replaced by pressure valves which open when the steam pressure, or in general the gas pressure, exceeds a certain level, so as to reduce the pressure, precisely as the *flyballs* functioned.

A simple modification transforms the automatic control system in Fig. 3.13 into a water clock, one of the simplest automatic systems for measuring the flow of time. A straw introduces regular flows of water, or even single drops, into a container or tank (even a large one) that can operate for many hours. The floating ball flows vertically, while attached to it is an indicator; on the wall of the container are notches indicating units of time. The gradual filling of the container, which is assumed to be regular, raises the floating ball to increasingly higher notches, and this measures the flow of time in the unit chosen for the notches: 15, 30, 60 min, etc., according to the experience of the clockmaker. When the container is completely filled a control mechanism sometimes operates in a manner opposite to that in Fig. 3.13. When the floating ball reaches its maximum level a butterfly valve is pushed that quickly empties the container, thus permitting the clock to begin operating again (this is a typical CI-PO impulse system, which will be illustrated in Sect. 4.3). Hourglasses function similarly: the upper part releases the sand to the lower part; when the latter is full, the manager turns over the hourglass and the cycle starts again.

A variant which simplifies the preceding mechanism (by eliminating the floating ball) consists simply in constructing a container—preferably of glass and cylinder shaped, with a narrow base and sufficient height—whose walls have regularly distanced notches with a predetermined time scale. The container is filled with water which is then let out through a small valve at its base. This regular outflow lowers the level of the liquid to increasingly lower notches, which allows the time flow to be measured and the time to be known by simply observing the notch corresponding to the water level. Though simpler than the preceding example, this clock has the disadvantage that the water outflow is never regular, since the pressure of the remaining liquid varies as the water flows out.

Even Galileo Galilei invented a similar mechanism, equally simple though not entirely automatic. Instead of filling up a graduated container, his design involved filling a container, of whatever form, and having the water regularly flow out through a straw at the bottom. The water would be collected in a second container at the base of the first and then weighed when it was time to measure how much time had passed. The weight is, in effect, converted into time using a simple scale derived from experience: for example, a kilogram of water could correspond to 10 min or an hour in relation to the amount of outflow in the upper container. If the weight-to-time conversion scales are accurate, this measure not only will provide more precise results but also have the advantage of permitting the calculation of the

submultiples: if a kilogram of water equaled 1 h, then 100 g of water would be equivalent to 6 min, etc.

I do not wish to always backtrack, but it is important to note that the first modern clocks to precisely measure time flow were the *pendulum clocks*, which function in a similar fashion to the flyballs in the system in Fig. 3.12, not to the floating ball in the model in Fig. 3.13, followed by the water clocks. We have to go back to Galileo, in the seventeenth century, to find the isochronism of the mechanical pendulum. The simplest mechanism to measure time is by a weight, a sphere, or a cylinder placed at the lower end of a rigid shaft with a pivot at the top end. A spring, or some similar device, moves the oscillating shaft while an instrument measures the number of oscillations, each of which is accomplished in the same period of time. The clockmaker measures the time for an oscillation based on his models, thus producing a unit of time; the duration of a period of time can thus be determined by multiplying the unit of time by the number of oscillations detected. If the time of an oscillation were 4 s, then 30 oscillations would measure 2 min. The addition of a quadrant provided the finishing touch for the pendulum clock we are all familiar with. The functioning similarity to the other mechanisms we have examined is clear. In Watt's *Centrifugal Regulator* the flyballs were moved by the centrifugal force generated by the rotating speed of the driveshaft; in the pendulum clock the weight was triggered by the force of gravity.

We can immediately observe that all clocks, water, sand, or pendulums, as well as modern wristwatches, are not control systems but "machines"; the variable that measures time flows—the height of the floating ball, the level and number of oscillations, or another internal "engine"—is not a variable that is controlled but simply the variable Y_t produced by the effector-clock. This is why all clocks, old as well as modern ones, must be periodically checked to eliminate the error, $E(Y)$, between the actual time (Y_t) and the time of reference (Y^*). The lever X on our clock (depending on the model) allows us to become the manager of the *manual* control system of time. For many digital clocks this *manual* control system can become entirely *automatic* if the clock has an internal detector apparatus to capture the international standard time or the UTC/GMT (Coordinated Universal Time), calculated using a weighted average of signals from a certain number of atomic clocks, located in several countries.

3.9.2   Halt Control Systems

"One-way" control systems are very common, whether or not the objective is in the form of a constraint that cannot be exceeded, a limit that is insurmountable, or a variable objective, as we typically see in systems referred to as "constraint-keeping" control systems. Figure 3.14 shows the model of a control system that allows us to park our car without crashing into the front wall of our garage.

"One-way" control systems are also common in our simplest daily activities and more numerous than steering control systems. It is true that, when we measure the quantity of rice for a meal for four people, we can calculate the right amount by
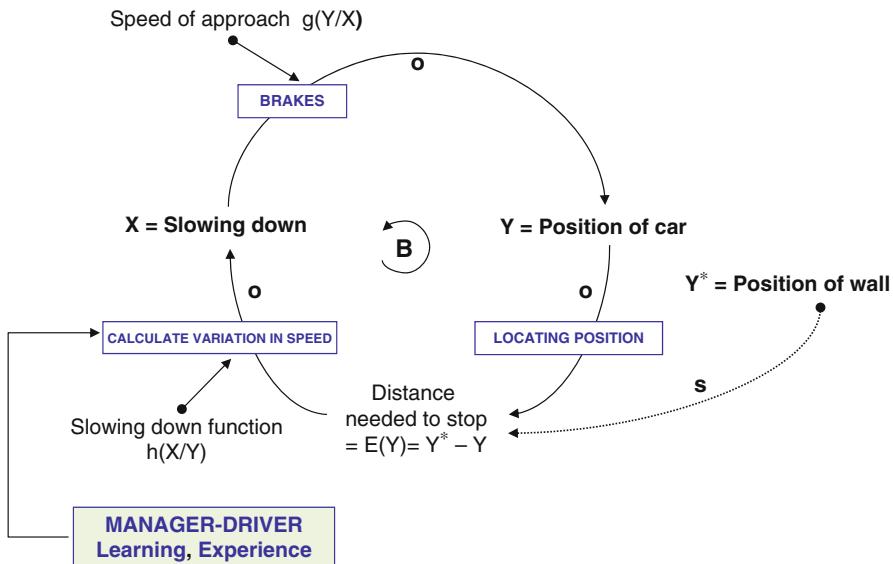
Fig. 3.14   Halt control system. Parking against a wall

incremental adjustments (up or down) in the rice before adding it to the water. However, when the rice is cooking and we wish to add the right amount of salt, it is better to do so in small, gradual quantities, since if we add too much the rice will become inedible. The rice also has to be cooked just right, and it must be tasted (detect the error) several times in order not to be overcooked.

In all mixtures and blends with several components (risotto, for example) the right amount of each component can never exceed the limit of the other components otherwise the product must be definitively eliminated.

It is not even necessary to give overly complicated examples. Figure 3.15 represents the control system that enables a *sommelier* to pour champagne always to the same (approximate) level in flute glasses.

3.9.3   Biometric Systems of Recognition and Identification

Going back to the distinction between systems of *attainment* and recognition (Sect. 3.2), let us consider *biometric systems*; that is, artificial *systems of recognition-identification* of a person based on several biological characteristics, which can be divided into physiological and behavioral aspects.

The physiological features refer to data that differentiates individuals in a singular way (static data), such as fingerprints, the shape of the hand, the form of the iris, the network of veins, or even the facial image. The behavioral characteristics seek to identify an individual based on the way the individual carries out at action, such as his handwriting, his style of walking, vocal features, and in general the way he performs an activity. Among the highly precise automatic systems of
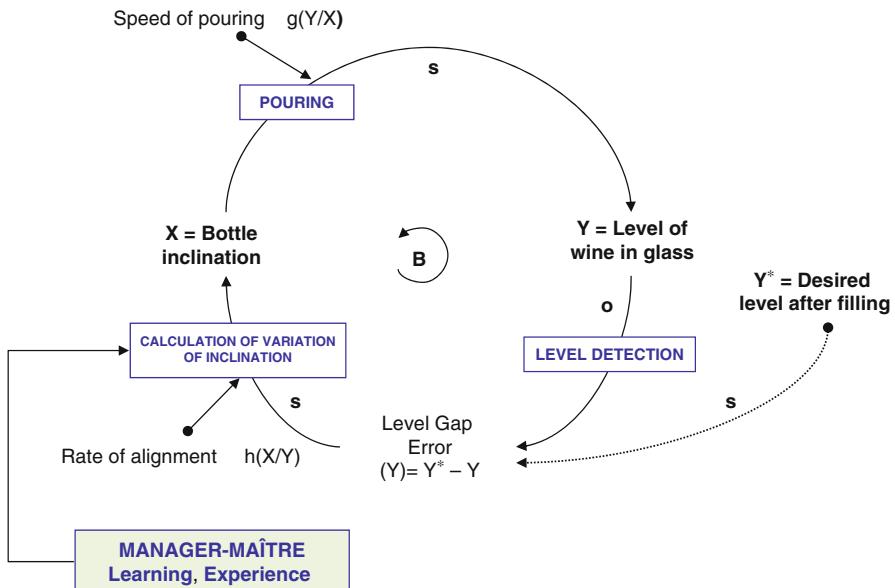
Fig. 3.15 Halt control system. Pouring champagne to the same level

recognition and identification are the systems of recognition of *fingerprints* from a continually updated archive, as shown in Fig. 3.16, which is derived from Fig. 3.3.

The fingerprint to recognize is obtained from an optical scanner and set as the objective (*Y**). The effector apparatus scans the archived fingerprints and the detector compares these and identifies the error (no match found) or the absence of error (match found).

The system of fingerprint recognition used today by the police is automatic (AFIS, Automated Fingerprint Identification System) and the high speed of recognition allows authorities to produce very large archives that can be quickly updated. In order to speed up the recognition of a fingerprint, AFIS uses classification algorithms of the fingerprint pattern and of the sampling of parts of the fingerprint itself or of systems of points. In addition to AFIS, fingerprint recognition is also widespread in many personal electronic products, ranging from laptop computers to cell phones, and is replacing recognition by means of badges. Another recognition system, similar to AFIS, is *hand* recognition; an appropriately arranged scanner obtains an image of the hand using an infrared camera (which captures the image in black and white, ignoring surface details) and sends the scanned profile of the hand to the sensor apparatus, which compares it to the image from a previously assembled archive.

The high speed and reliability of recognition systems has made possible more ambitious *pattern recognition* techniques applied to the face or physique of individuals. Such systems are based on the fact that the face (or body structure) of a person can be recognized not by observing it in all its details but by referring to
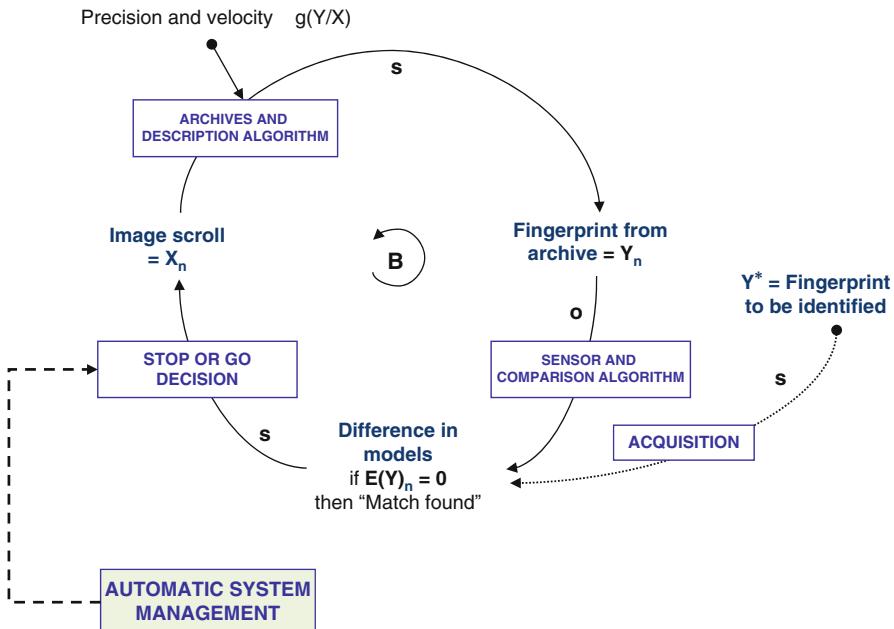
Fig. 3.16   Model of automated fingerprint identification system

certain fundamental characteristics; for example, the color of the eyes or the distance between them; the shape and size of the nose and ears; the broadness of the forehead; the shape of the chin; hair color and styling.

An automatic *pattern recognition* system for the face is a control system that acquires images from some scanning system and identifies the proportions, distances, and ratios between the various somatic features. It forms a model that represents the objective ($Y*$) to recognize and compares this with the models in the archive. Systems that recognize people by their retina or iris are very accurate. We have all seen machines that scan the retina to identify a person in order to authorize access to some device.

A DNA recognition system was created at the start of the 1990s involving data banks containing DNA profiles which, under the law, can be accessed and preserved, for example, the DNA of people who have committed certain types of crimes and those who have had restrictions placed on their personal freedoms; DNA samples from biological evidence acquired during criminal proceedings; the DNA of missing people or their relatives and of unidentified corpses or body remains. A DNA sample taken at some site is analyzed and the profile is compared with those in the data base; if recognized, then the person it belongs to is identified. Not all countries have DNA data banks, which is understandable given the delicacy of the problem of the privacy protection of citizens and of the safekeeping of samples and access to them.

This type of control system functions just like that in the model in Fig. 3.16.

3.9.4   Explorative Systems

The systems of recognition presented in Sect. 3.2 also allow us to describe the functioning of *explorative systems*. Exploring an unknown environment of whatever kind, from a territory to the web, a bookstore to a just-purchased book, from craters on Mars to the starry firmament, signifies, in fact, scanning the objects of this environment and comparing these with an archive of models. The more the objects in the environment are identified, the better known the explored environment becomes. Since they are not in the archive, the unidentified objects are transformed into models (description, designs, photos, films, etc.) and then *included* in the archive, which gradually grows to make future explorations easier.

In order to understand the exploration process as an activity of a control system of *recognition*, let us imagine we have a dark room of unknown content which we are forced to explore; we move around without being able to see anything, only perceiving tactile sensations, for example, when we run into something, or sounds connected to these impacts. Our mind tries to identify the "objects" that could have caused these perceptions, and we can thus identify a sharp-cornered piece of furniture or the sharp edge itself, a glass that has fallen and shattered, a soft cushion, a carpet, etc.

The Center for Space Observation is exploring a swath of the night sky and "strange signals" arrive from space, which are captured by instruments. The astronomers immediately analyze these signals and look in their "atlas of celestial bodies" to identify objects that could be the source of the signals. Let us assume we are able to identify a pulsar located in the *XYZ* coordinates, which seems to be the source. The exploration has then finished. In the event no catalogued "object" can be linked to those signals, the astronomers would presumably add a report to their catalogue: "Today at MM hours signals were identified with the following characteristics (position, mass, waves emitted, etc.) which we have named "H3b12." From that moment on, any astronomer identifying similar signals can use the catalogue to recognize the sender of the signal.

If we substitute "astronomer" with "doctor" and "strange signals from space" with "strange symptoms of patient," you will find an explanation for the many papers found in medical journals. Likewise, if we substitute "astronomer" with "nature explorer in new lands" and "strange signals from space" with "unknown fauna and flora" we will find ourselves going back to Charles Darwin's *Voyage of the Beagle*. We have represented the *explorative* activity as the system of recognition and identification shown in Fig. 3.3, including, in the box at the lower right, the "instructions for updating the catalogue"; for example, in the following way:

$$\text{for } 1 \leq n \leq N, \text{if } Y_n \neq Y^*, \text{then } Y^* = Y_{N+1} \tag{3.1}$$

A variant of explorative systems are systems of *improvement*, characterized by an increase in the objective when the exploration identifies values for Y_n which are higher than those of the objective. Belonging to this type of system are
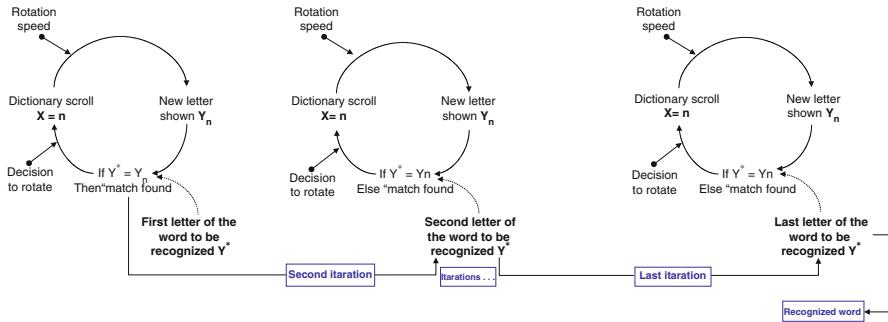
Fig. 3.17 Control system for recognizing a word in the dictionary by sequencing its component letters

selecting systems: an exploration of the environment leads to selecting the "best" object in terms of one of its qualities: the highest, sweetest, most flavorful, coldest, most stocked with merchandise, wealthiest, etc. This selection must occur through a comparison between the objects to be selected, which are identified by exploring the environment, and the temporary object considered to be the "best" and set as the objective. In order to transform a system of recognition and identification into a *selecting system* we must replace (3.1) in Fig. 3.3 by the following instruction:

$$\text{for } 1 \leq n \leq N, \text{if } Y_n > Y^*, \text{then } Y^* = Y_n. \tag{3.2}$$

This instruction differs from (3.1) in that it produces a continual increase in the objective in selecting other "best" objects.

3.9.5 Looking Up a Word in the Dictionary

The control system for looking up the word Y^* (Sect. 2.16.3) can operate by directly comparing the entire word with every word in the dictionary using the model in Fig. 2.21; however, as shown in Fig. 3.17, in order to do so it must read the first letter of Y^* and start scanning through the dictionary, page after page, until it comes to the words whose first letter is the same as that in Y^*.

It then reads the second letter and continues scanning until it arrives at the words whose second letters are the same as in Y^*, proceeding in this manner until it arrives at the last letter in Y^*. If the dictionary contains a word whose last letter is the same as the last letter of Y^*, then the search has ended and we can read the definition or translation.

3.9.6 Achilles and the Tortoise: Zeno's Paradox

Aristotle writes that, in order to support the idea that reality is made up of a single and immutable Being, Zeno of Elea proposed several paradoxes to show, absurdly, the illusion of both plurality and motion, despite what our daily lives clearly reveal to us.
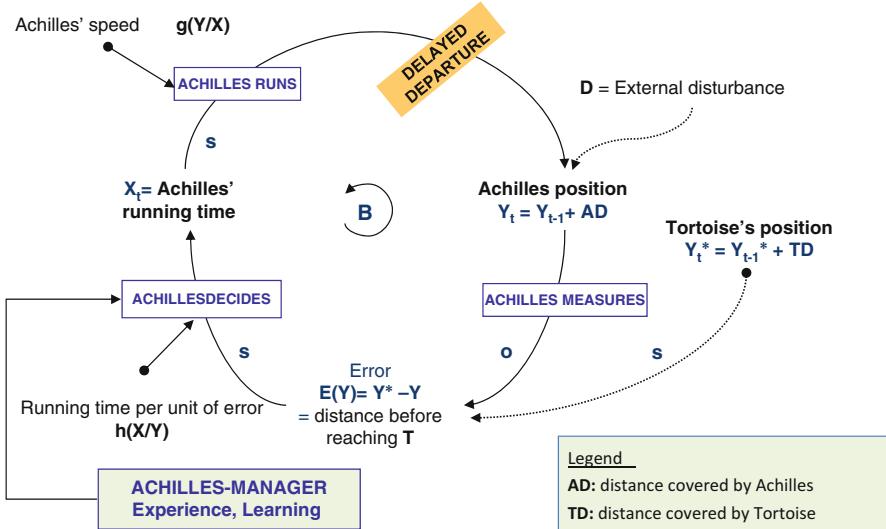
Fig. 3.18   Control system which Zeno had in mind. Achilles fails to pass the turtle

The most famous of Zeno's paradoxes is probably that which argues against motion, whereby the philosopher wants to show that, despite the fact Achilles was faster than a tortoise, he could never have caught up to and passed it in a race if the latter had had a head start (the paradox is summarized by Smith (2013), in Prime). As we can see in the model in Fig. 3.18, the action variable, X_t, represents the amount of time Achilles decides to run after having observed the position of the tortoise at time "t"; the variable Y_t, on the other hand, is his position after having run the amount of time decided at the previous instant.

In fact, starting after an amount of time, let us say $t = n$, Achilles, in order to overtake the tortoise, would have had to arrive at where the latter was positioned at "n"; however, in the meantime the tortoise advances an additional, shorter distance, which Achilles must bridge. When Achilles reaches this new position the tortoise advances again some distance; this continuous pursuit translates into an infinite series of small mutual advancements, increasingly shorter in distance, so that from this perspective the gap between Achilles and the tortoise could never have been eliminated.

This is a classic example control systems of *pursuit*; in this case Achilles, the system's manager, has the objective of reaching the observed position of the tortoise (Y_t^*) at every "t" by changing his position by means of decisions to run for a sufficient length of time to eliminate the gap.

Figure 3.19—even though it involves a somewhat reduced number of control cycles and of decimal places—shows that Zeno's example calls for its effects to occur punctually. However, this occurs because the system in Fig. 3.18 (as conceived of by Zeno) *is poorly designed*.

			DYNAMIC OBJECTIVE	
ATR = E/AS	AP = ATR *AS	DST = ATR * TS	$TP_t* = TP_{t-1} + DST$	E = TP - AP

TS TORTOISE'S SPEED 5
AS ACHILLES' SPEED 20

STEP	Achilles' time running ; DELAY	Achilles' position	Delta space, tortoise	Tortoise's position	E = Gap to eliminate
0	1.00	0.00	5.00	5.00	5.00
1	0.25	5.00	1.25	6.25	1.25
2	0.0625	6.2500	0.3125	6.5625	0.3125
3	0.01562500000000000000	6.56250000000000000000	0.07812500000000000000	6.64062500000000000000	0.07812500000000000000
4	0.00390625000000000000	6.64062500000000000000	0.01953125000000000000	6.66015625000000000000	0.01953125000000000000
5	0.00097656250000000000	6.66015625000000000000	0.00488281250000000000	6.66503906250000000000	0.00488281250000000000
6	0.00024414062500000000	6.66503906250000000000	0.00122070312500000000	6.66625976562500000000	0.00122070312500000000
7	0.00006103515625000000	6.66625976562500000000	0.00030517578125000000	6.66656494140625000000	0.00030517578125000000
8	0.00001525878906250000	6.66656494140625000000	0.00007629394531250000	6.66664123535156000000	0.00007629394531250000
9	0.00000381469726562500	6.66664123535156000000	0.00001907348632812500	6.66666030883789000000	0.00001907348632812500
10	0.00000095367431640625	6.66666030883789000000	0.00000476837158203125	6.66666507720947000000	0.00000476837158203125
11	0.00000023841857910156	6.66666507720947000000	0.00000119209289550781	6.66666626930236000000	0.00000119209289550781
12	0.00000005960464477539	6.66666626930236000000	0.00000029802322387695	6.66666656732559000000	0.00000029802322387695
13	0.00000001490116119385	6.66666656732559000000	0.00000007450580596924	6.66666664183139000000	0.00000007450580596924
14	0.00000000372529029846	6.66666664183139000000	0.00000001862645149231	6.66666666045784000000	0.00000001862645149231
15	0.00000000093132257462	6.66666666045784000000	0.00000000465661287308	6.66666666511446000000	0.00000000465661287308
16	0.00000000023283064365	6.66666666511446000000	0.00000000116415321827	6.66666666627861000000	0.00000000116415321827
17	0.00000000005820766091	6.66666666627861000000	0.00000000029103830457	6.66666666656965000000	0.00000000029103830457
18	0.00000000001455191523				
19	0.00000000000363797881				
20	0.00000000000090949470				
21	0.00000000000022737368				
22	0.00000000000005684342				
23	0.00000000000001421085				
24	0.00000000000000355271				
25	0.00000000000000088818				
26	0.00000000000000000000				
27	0.00000000000000000000				
28	0.00000000000000000000				
29	0.00000000000000000000				
30	0.00000000000000000000				
31	0.00000000000000000000				
32	0.00000000000000000000				
33	0.00000000000000000000				
34	0.00000000000000000000				
35	0.00000000000000000000				
36	0.00000000000000000000				
37	0.00000000000000000000				
38	0.00000000000000000000				
39	0.00000000000000000000				
40	0.00000000000000000000				
41	0.00000000000000000000				
42	0.00000000000000000000	6.66666666666666000000	0.00000000000000000000	6.66666666666667000000	0.00000000000000000000

Fig. 3.19 A simulation of Zeno's paradox

In order to modify the regulation decisions Achilles must rationally make, we need only transform the control system from a *variable objective* (the tortoise who is fleeing) into a *fixed objective* (the *finish line*) system.

In this new system Achilles does not have to decide the *running time* needed to catch up to the tortoise but the *time needed to reach the finish line*, as illustrated in the model in Fig. 3.20. As we can easily see from the calculations, it is guaranteed that Achilles will surpass the tortoise, on the condition that his speed is greater than his adversary's and there is a sufficient distance to the finish line.

In fact, what is truly different in this second system with respect to the previous one is the procedure for quantifying the error: in the first case this is calculated with respect to the *variable objective* of the tortoise's position, while in the second in terms of the *fixed finish line* (objective).

3.9.7 Variable-Objective Systems: Malthusian Dynamics

In Sect. 3.4 I have defined as "variable-objective systems" those control systems whose objective, *Y**, changes over time, so that the system appears to follow a continually changing objective.

	TS	TORTOISE'S SPEED		10	Y* = Distance to the		500	
	AS	ACHILLES' SPEED		20	finish line			
		DELAY				FIXED OBJECTIVE		
		AD = tAS	TD = t TS		GAP = DT - DA	Y*		E = Y* - AD

STEP	DISTANCE TRAVELED BY ACHILLES / DELAY	DISTANCE TRAVELED BY TORTOISE	Gap	Finish line	E = ACHILLES' GAP TO FINISH LINE
0	0.0	0.0	0.0	500.0	500.0
1	0.0	10.0	-10.0	500.0	500.0
2	0.0	20.0	-20.0	500.0	500.0
3	0.0	30.0	-30.0	500.0	500.0
4	0.0	40.0	-40.0	500.0	500.0
5	0.0	50.0	-50.0	500.0	500.0
6	20.0	60.0	-40.0	500.0	480.0
7	40.0	70.0	-30.0	500.0	460.0
8	60.0	80.0	-20.0	500.0	440.0
9	80.0	90.0	-10.0	500.0	420.0
10	100.0	100.0	0.0	500.0	400.0
11	120.0	110.0	10.0	500.0	380.0
12	140.0				360.0
13	160.0				340.0
14	180.0				320.0
15	200.0				300.0
16	220.0				280.0
17	240.0				260.0
18	260.0				240.0
19	280.0				220.0
20	300.0				200.0
21	320.0				180.0
22	340.0				160.0
23	360.0				140.0
24	380.0				120.0
25	400.0				100.0
26	420.0				80.0
27	440.0				60.0
28	460.0				40.0
29	480.0	290.0	190.0	500.0	20.0
30	500.0	300.0	200.0	500.0	0.0
31	520.0	310.0	210.0	500.0	-20.0

Fig. 3.20 Control System Achilles employs. Achilles surpasses the tortoise

Figure 3.21 presents a typical example of a *system of pursuit* where there is a population dynamics—whose initial size (equal to 1,000 units) is specified—which varies annually dues to births and deaths. Since the birth rate is assumed to be greater than the death rate, year-to-year there is a gradual increase in population. However, there is also assumed to be a limited quantity of resources (for convenience sake, food supplies that each year increase by the same amount; that is, food is not a nonrenewable resource). Initially we assume that the amount of resources (7,500 units) is more than what is necessary for consumption, which is 5,000 units. Thus, the population can expand until the limit imposed by the availability of limited resources, which are completely used up at $t = 9$.

The simulation program also allows for the possibility that the population, having reached the resource saturation point, can initiate a *production process* to increase the resources and continue to expand; however, this growth will not continue for long since we assume there is a limit (10,000 units) to the expansion of the scarce resource due, for example, to the constraint from the need for other resources for its production (land, water, fertilizer, etc.). Therefore, the population expansion stops at $t = 17$. In short, with the data inserted in the control panel, the population will expand until $t = 9$ with the available natural resources, at which

						reaction time	**1**				
initial population			**1000**	production			**5.0**				
birth rate			**10%**	if famine			**9%**	negative percentage			
mortality rate			**5%**	if famine			**6%**	negative percentage			
per capita food cons.			**5.0**	if famine			**5.0**	negative percentage			
food as resource			**7500**	maximum food			**10000**				

years	old populat. + births – deaths	birth rate	births	mortality rate	deaths	unit food consumption	food requirement	food available	humanitarian aid	productivity	production	food surplus /deficit
0	1000.00	10.0%	100	5.0%	50	5.0	5000	7500		5.0	0	-2500
1	1050.00	10.0%	105	5.0%	53	5.0	5250	7500		5.0	0	-2250
2	1102.50	10.0%	110	5.0%	55	5.0	5513	7500		5.0	0	-1988
3	1157.63	10.0%	116	5.0%	58	5.0	5788	7500		5.0	0	-1712
4	1215.51	10.0%	122	5.0%	61	5.0	6078	7500		5.0	0	-1422
5	1276.28	10.0%	128	5.0%	64	5.0	6381	7500		5.0	0	-1119
6	1340.10	10.0%	134	5.0%	67	5.0	6700	7500		5.0	0	-800
7	1407.10	10.0%	141	5.0%	70	5.0	7036	7500		5.0	0	-464
8	1477.46	10.0%	148	5.0%	74	5.0	7387	7500		5.0	0	-113
9	1551.33	10.0%	155	5.0%	78	5.0	7757	7757		5.0	257	0
10	1628.89	9.0%	147	5.0%	81	5.0	8144	8144		5.0	644	0
11	1694.05	9.0%	152	5.0%	85	5.0	8470	8470		5.0	970	0
12	1761.81	9.0%	159	5.0%	88	5.0	8809	8809		5.0	1309	0
13	1832.28	9.0%	165	5.0%	92	5.0	9161	9161		5.0	1661	0
14	1905.58	9.0%	172	5.0%	95	5.0	9528	9528		5.0	2028	0
15	1981.80	9.0%	178	5.0%	99	5.0	9909	9909		5.0	2409	0
16	2061.07	9.0%	185	5.0%	164	5.0	10305	10000		5.0	2500	305
17	2082.44	9.0%	187	6.0%	207	5.0	10412	10000		5.0	2500	412
18	2062.47	9.0%	186	6.0%	186	5.0	10312	10000		5.0	2500	312
19	2061.87	9.0%	186	6.0%	186	5.0	10309	10000		5.0	2500	309
20												309
21												309
22												309
23												309
24												309
25												309
26												309
27												309
28												309
29												309
30												309
31												309
32												309
33												309
34												309
35												309
36												309
37												309
38												309
39												309

ECOSYSTEM WITH MALTHUSIAN DYNAMICS and scarce non-replenishable resources

Fig. 3.21 Ecosystem with Malthusian dynamics and scarce renewable resources

point a production process for the resource will start, increasing the latter to 10,000 units. Subsequently, from $t = 17$ the population will stabilize at 2,062 units.

This simple simulation program allows for other hypotheses; in particular, that once the saturation point of the resource occurs there could be a change in the birth rate (downward) or the death rate (upward), in addition to the hypothesis of unit consumption (rationing). Thus Fig. 3.21 represents a variable-objective system; in this case, the objective is represented by the admissible population, taking account of the scarce resource, whose limit, however, varies due to new production until the saturation point is reached. Figure 3.22 simulates the case where the process

reaction time	1		
initial population	1000	production	5.0
births rate	10%	if famine	10% negative percentage
deaths rate	5%	if famine	5% negative percentage
per capita food cons.	5.0	if famine	5.0 negative percentage
food as resource	7500		
cons. of resource	1	maximum depletable resource	15000 non-replenishable

years	old population + births – deaths	birth rate	births	mortality rate	deaths	unit food consumption	food requirement	food available	humanitarian aid	productivity	production	surplus or deficit	consumption of resource	resource available	maximum production
0	1000	10%	100	5.0%	50	5.0	5000	7500		5.0	0	-2500	-	15000	15000
1	1050	10%	105	5.0%	53	5.0	5250	7500		5.0	0	-2250	-	15000	15000
2	1103	10%	110	5.0%	55	5.0	5513	7500		5.0	0	-1988	-	15000	15000
3	1158	10%	116	5.0%	58	5.0	5788	7500		5.0	0	-1712	-	15000	15000
4	1216	10%	122	5.0%	61	5.0	6078	7500		5.0	0	-1422	-	15000	15000
5	1276	10%	128	5.0%	64	5.0	6381	7500		5.0	0	-1119	-	15000	15000
6	1340	10%	134	5.0%	67	5.0	6700	7500		5.0	0	-800	-	15000	15000
7	1407	10%	141	5.0%	70	5.0	7036	7500		5.0	0	-464	-	15000	15000
8	1477	10%	148	5.0%	74	5.0	7387	7500		5.0	0	-113	-	15000	15000
9	1551	10%	155	5.0%	129	5.0	7757	7500		5.0	257	257	257	14743	14743
10	1578	10%	158	5.0%	105	5.0	7888	7757		5.0	388	131	388	14356	14356
11	1630	10%	163	5.0%	134	5.0	8151	7888		5.0	651	263	651	13704	13704
12	1659	10%	166	5.0%	112	5.0	8295	8151		5.0	795	144	795	12909	12909
13	1713	10%	171	5.0%	140	5.0	8566	8295		5.0	1066	270	1066	11843	11843
14	1745	10%	174	5.0%	119	5.0	8724	8566		5.0	1224	158	1224	10620	10620
15	1800	10%	180	5.0%	146	5.0	9002	8724		5.0	1502	278	1502	9118	9118
16	1835	10%	183	5.0%	126	5.0	9174	9002		5.0	1674	172	1674	7444	7444
17	1892	10%	189	5.0%	152	5.0	9461	9174		5.0	1961	287	1961	5483	5483
18	1929	10%	193	5.0%	134	5.0	9647	9461		5.0	2147	186	2147	3336	3336
19	1989	10%	199	5.0%	159	5.0	9943	9647		5.0	2443	296	2443	893	893
20	2029	10%	203	5.0%	142	5.0	10144	9943		5.0	893	201	893	-	-
21	2090	10%	209	5.0%	516	5.0	10450	8393		5.0	0	2057	-	-	-
22	1783	10%	178	5.0%											
23	1589	10%	159	5.0%											
24	1579	10%	158	5.0%											
25	1579	10%	158	5.0%											
26	1579	10%	158	5.0%											
27	1579	10%	158	5.0%											
28	1658	10%	166	5.0%											
29	1741	10%	174	5.0%											
30	1787	10%	179	5.0%											
31	1789	10%	179	5.0%											
32	1789	10%	179	5.0%											
33	1589	10%	159	5.0%											
34	1579	10%	158	5.0%											
35	1579	10%	158	5.0%											
36	1579	10%	158	5.0%											
37	1579	10%	158	5.0%											
38	1579	10%	158	5.0%											
39	1579	10%	158	5.0%											
40	1579	10%	158	5.0%											
41	1579	10%	158	5.0%											
42	1579	10%	158	5.0%											
43	1579	10%	158	5.0%											
44	1579	10%	158	5.0%											
45	1579	10%	158	5.0%											
46	1579	10%	158	5.0%											
47	1579	10%	158	5.0%											
48	1579	10%	158	5.0%											
49	1579	10%	158	5.0%											
50	1579	10%	158	5.0%											

ECOSYSTEM WITH MALTHUSIAN DYNAMICS with limited resources

(chart, vertical axis scaled 0 to 12000; horizontal axis 0 to 60; curves labelled: food requirement, food limit, population)

Fig. 3.22 Ecosystem with Malthusian dynamics and scarce nonrenewable resources

depends on the use of a *nonrenewable* productive resource available only in limited quantities (15,000 units).

If there were no limits to obtaining the resource, the population would continue to increase while producing more and more resources, in an endless pursuit. By posting a limit to the reproducible resource the system tends to the above limit gradually but inevitably. The production starts at $t = 9$, when the food available in the system, which can be reproduced year after year (in the amount of 7,500 units), is completely utilized. Food is produced above 7,500 units gradually until t = 19, when the scarce resource is completely utilized (last three columns of the table in Fig. 3.22). From then on the production of food ceases and the population quickly falls to 1,579 units ($t = 24$).
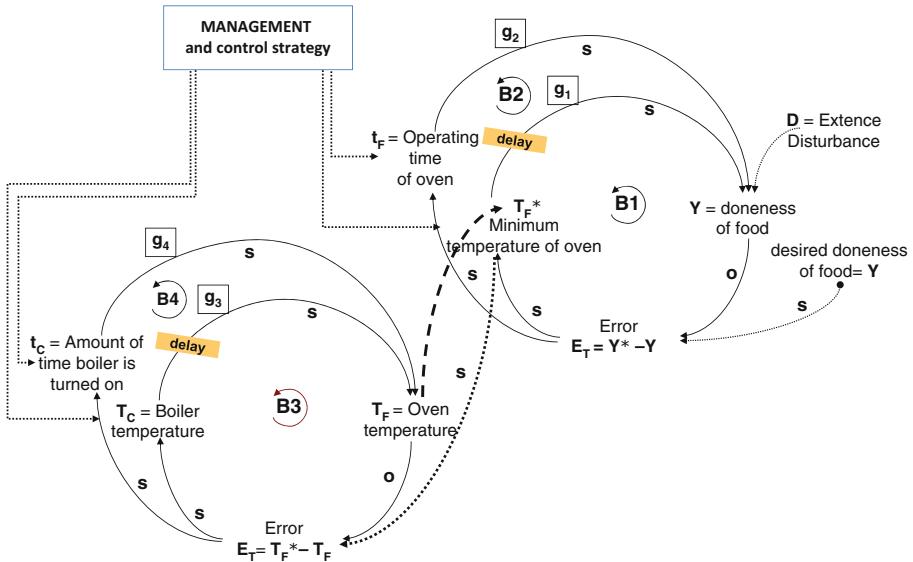
Fig. 3.23 Serial systems. The oven and boiler

Does not the possibility exist of humanitarian aid? Of course! However, the simulation shows that such aid—which would begin at $t = 27$—would have, being an "external" disturbance, a short-lived effect; when this aid ceases the system inevitably *controls the population*, which, at $t = 33$, returns to the *maximum allowed limit* of 1,579 units. This second example highlights, though in a simplified manner, the fact that population growth, and the consequent growth in production, must take into account the constraints imposed by the ecosystem's sustainability.

3.9.8   Serial Systems: The Oven and the Boiler

Figure 3.23 represents a control system for how long food is cooked in an oven that can be regulated in terms of temperature and cooking time.

The oven temperature is, in turn, a variable controlled by a second control system represented by the boiler, which provides the oven temperature through two levers: the boiler temperature and the amount of time the boiler is turned on. It is not possible to cook food by lengthening the operating time of the oven, tF, if the oven temperature, TF, is too low; thus, in order to cook food we need a minimum oven temperature, TF^*, which is set as the objective of the second system (the boiler) which, through its two levers, allows the oven temperature, TF, to reach the minimum necessary to cook the food.

3.9.9   Qualitative Control Systems: Procedure to Determine the States of Variety of the Qualitative Variables

In order for the *qualitative* control systems in Sect. 3.2 to function, it is necessary to identify or define appropriate "*scales of variety*," or *qualitative scales*, for the qualitative dimensions of color, flavor, smell, tactile sensation, etc., which enables a *qualitative state* of the variable to control, Y_t, to be compared with a *qualitative state* set as the objective, Y^*, as we clearly saw in the example of recognizing the replacement button for our clothing or the wine at a wine-tasting competition. The determination of the "states of variety" is not easy since, in undertaking the qualitative determination, it is not possible to apply procedures of enumeration, measurement, and metricization. We start by distinguishing between *direct* and *indirect* qualitative determination; in the first case the process to determine each qualitative dimension specified several elementary states, while indirect determination derives the states using operations carried out on elementary states.

Specifying the name, order of finish, license plate, or identity card number falls under direct determination. Specifying the state of purity of a diamond instead comes under indirect determination, since the state of purity is obtained by, for example, combining the quality of the color, the degree of impurity, and the transparency of the diamond. While the color of a button is determined in a direct manner based on "color scales," the flavor of a wine represents a mix of a number of qualitative states which the wine taster is able to recognize.

Among the possible determination procedures, I will only mention here the following, all of which are general and widely applicable:

- Specific procedure
- Analogic procedure
- Diagnostic procedure
- Defining or conventional procedure

The *specific procedure* searches for the qualitative state of the dimension to determine through observations or specific research. It is applied, for example, in attributing and searching for names or the specific signs of the objects of observation, in identifying their composition, specifying their structure, and so on. This procedure is not simple at all. Consider the difficulties in identifying the name of a crime suspect who refuses to give information about himself, or in identifying the structure of a biological system (cellular, the human brain) or the composition of a chemical mix or compound.

The *analogic procedure* involves specifying the quality of a given object or phenomena by comparing this to a predetermined and known scale of varieties. This is the preferred procedure for qualifying colors, flavors, odors, noises, and so on. We have all learned the simplest scales of variety through education, which we refer every day. We say our car is red, that we have tasted food with a prevalently sour flavor, or that we have bought lavender soap. When we wish to be more precise we may specify that our car is "Ferrari red" or "blood red" and that the food tastes of

bitter lemon or of lemon mixed with orange. The analogic procedure makes the determinations more precise by using a *standard scale* (chromatic, balance of flavors, olfactory, acoustic, etc., recognized in the context of a group of operators) which is referred to in order to specify, with the maximum precision that is possible, admissible and practical, to which element in the scale the dimension to determine can be analogically compared. The difficulties in doing so are well known to manufacturers of colors, paints, colored fabric, restorers, cooks, wine producers, perfume manufacturers, musicians, piano tuners, and so on.

A particular type of analogic procedure is undertaking a *comparison with a sample or prototype,*" which is applied to specify the quality of complex objects (facial shape of a person, esthetic aspect of an object, structure, etc.) through a "line by line" comparison with a sample object, which represents the only state of comparison of a hypothetical scale of reference and which is set as the objective of the qualitative control system.

The *diagnostic procedure* occurs in all cases in which the "way of being" of a qualitative dimension is spelled out by a definition that indicates the element or elements characterizing the various states of the scale of variety. In order to specify the "way of being" of a qualitative dimension, specific surveys can be undertaken to identify "symptoms" or "signs" that can distinguish between the various states of the dimension. This procedure is adopted to identify the state of illness of a given individual or the state of insolvency of an enterprise. The diagnostic procedure is also widely used in the judicial sphere. In fact, many norms contain definitions of "status" (parent, owner, partner, firm, etc.) or define particular types of action (declaring, registering, informing, adopting, selling, paying off debt, etc.).

In many cases the qualitative dimensions appear so detailed and complex that their quality is conventionally decreed through appropriate *circumscribing,* or *conventional definitions.* In this case the determination of the "states of variety" of a dimension consist in verifying whether or not the determination conforms to the conventional definition. Consider how difficult it still is to ascertain whether or not a biological entity possesses the quality of "belonging to the living sphere." When is a biological structure deemed to be living (animated)? How can an exploratory system—a space probe on a new planet—recognize when it encounters a living biological being? Biologists have provided circumscribed definitions for the notion of "living," definitions that change along with scientific and technical progress which increase our capacity to understand phenomena and refine our observations.

The "macroscopic living being" (perceivable to the naked eye or by using modest magnifying lenses) gave way to Galileo's invention of the microscope that enabled man to recognize the "microscopic living being": single-cell organisms, bacteria, viruses, and so on. This made it necessary to deal with the problem of their inclusion in the sphere of the living (animated) being, which entailed modifying or refining the definition of "living." Equally as complex is ascertaining the quality "belonging to the category '*alive.*'" When a biological structure belongs to the category of the living, is it *alive*? When does it cease being alive? This

determination is particularly difficult in pathological or "cutting-edge" cases we can all imagine.

The above considerations and examples reveal how the procedures for qualitative determinations are not always individually applicable: instead the methods of qualification must often be used in combination. Moreover, the methods of qualitative determination, especially the definition procedure, often refer to specific ordinal or cardinal scales that allow the states of the scales of variety to be more accurately determined.

3.10   Summary

Single-lever and single-objective control systems can be divided into several interesting categories, of which the following are particularly relevant:

(a) Artificial and Natural Control Systems (Sect. 3.1)
(b) Manually Controlled systems, Automatic Control Systems, or Cybernetic Systems (Sect. 3.1)
(c) Quantitative and Qualitative Control Systems (Sect. 3.2)
(d) Attainment and Recognizing Systems (Sect. 3.2)
(e) Steering and Halt Control Systems (Sect. 3.3)
(f) Fixed- and Variable-Objective Systems (or systems of "pursuit") (Sect. 3.4)
(g) Collision, Anti-collision, and Alignment Systems (Sect. 3.5)
(h) Tendential and Combinatory Control Systems (Sect. 3.6)
(i) Parallel or Serially Connected Systems (Sect. 3.7)
(j) Control Systems forming an Holarchy (Sect. 3.8)

Chapter 4
The Ring Completed: Multi-lever and Multi-objective Control Systems

There once were two watchmakers, named Hora and Tempus, who manufactured very fine watches. Both of them were highly regarded, and the phones in their workshops rang frequently—new customers were constantly calling them. However, Hora prospered, while Tempus became poorer and poorer and finally lost his shop. The watches the men made consisted of about 1,000 parts each. Tempus had so constructed his that if he had one partly assembled and had to put it down—to answer the phone, say—it immediately fell to pieces and had to be reassembled from the elements. [. . .] The watches that Hora made were no less complex than those of Tempus. But he had designed them so that he could put together subassemblies of about ten elements each. Ten of these subassemblies, again, could be put together into a larger subassembly; and a system of ten of the latter sub-assemblies constituted a whole watch. Hence, when Hora had to put down a partly assembled watch in order to answer the phone, he lost only a small part of his work, and he assembled his watches in only a fraction of the man-hours it took Tempus. (Herbert Simon 1962, p. 90).

Chapter 2 presented the general notion of a single-lever control system, illustrating the logical as well as technical structure of such a system. In addition, Chap. 3 proposed several important types of control systems which, though dissimilar in appearance, satisfy the definition and general logic of single-lever control systems. This chapter presents a broader view of control systems by introducing two important generalizations: (1) *multi-lever* control systems and (2) *multi-objective* control systems. Multi-lever control systems can have dependent or independent levers. Multi-objective control systems can be apparent or effective. Generalizing further, we will examine the notion of *impulse systems*, which play a fundamental role in life. With *multi-lever* systems it is fundamental to understand the concept of *control strategy*, that is, programming the activation of the various levers to achieve the objective. In *multi-objective* systems the choice of strategy is coupled with the definition of the *control policy*, which chooses the order of priorities regarding actions on the various objectives. Specifying the control strategies requires introducing the concept of *cost–benefit analysis* applied to the various levers. Specifying the control policies brings up the notion of *scale of priorities* regarding the various objectives.

4.1 Dual-Lever Control System with Mutually Dependent Levers

In order to make it easy to understand the general logic of control, the previous chapters presented *single-lever* control systems—where a single lever, X_t, is sufficient to direct the trajectory of the variable to control, Y_t, toward the objective Y^*—since these are relatively simple to illustrate, even though they often reveal complex dynamics.

It is time now to abandon the limitation of a single control lever and assume that the *passive* variable, Y_t, can be controlled through "N" *active* variables: $X1_t$, $X2_t, \ldots, Xn_t, \ldots XN_t$. We thus can form control systems with two, three, or multiple control "levers," which I shall define as *pluri-lever* or *multi-lever* systems.

As in single-lever systems, in multi-lever ones the values of the control levers depend on the amount of error $E(Y)$. Based on the manner in which the levers depend on $E(Y)$, we can divide multi-lever systems into two general categories: *monolayer* and *multilayer* systems.

To simplify the analysis of multi-lever control systems it is appropriate to first examine *monolayer* systems, postponing until Sect. 4.5 the examination of *multilayer* ones. I shall begin with the simplest case of the *dual-lever* control system, where Y_t (effect) is controlled by $X1_t$ and $X2_t$ (cause).

I define *monolayer* control systems as those where all the levers are regulated by the manager based on the "general," or *first-level*, error: $E^1(Y)_t = Y^* - Y_t$ (we assume that there is a fixed objective). This means that the system can be described "simply" by generalizing the recursive equations (2.6) to (2.8), which were introduced in Sect. 2.3. In order to simplify the notation (until the discussion of *multilayer* systems in Sect. 4.5), the first-level error will always be indicated by $E(Y)_t$ to avoid confusion.

Assuming that the *reaction times* are equal to "1" for both levers, the action rate "g" and reaction rate "h" are constant, $t = 0, 1, 2, \ldots$, and that there are no delays, then the new recursive equations for the *dual-lever* systems become

$$Y_n = [X1_n g_1 + X2_n g_2] + D_n \tag{4.1}$$

$$E(Y)_n = Y^* - Y_n \tag{4.2}$$

$$X1_{n+1} = X1_n + E(Y)_n h_1; \tag{4.3}$$

$$X2_{n+1} = X2_n + E(Y)_n h_2. \tag{4.4}$$

It is immediately clear from the above equations that the manager must carefully decide how to regulate the value of the levers; if he or she applies (4.3) and (4.4) in a cumulative manner (taking into account for both the amount of $E(Y)_n$) and also applies (4.1) with the new values for the levers, he or she could cause Y_{n+1} to become too large, thus making the control difficult, if not impossible.

Thus, the manager must decide in what *proportion* to regulate the two control levers by determining their *relative weights*, s_1 and s_2—$0 \leq s_1 \leq 1$, $0 \leq s_2 \leq 1$, $s_1 + s_2 = 1$—, in order to complete (4.1) as follows:

$$Y_n = [s_1(X1_n g_1) + s_2(X2_n g_2)] + D_n$$

Clearly, in order to control Y_t the manager must carry out a convex linear combination of the two levers. The vector [s_1 and s_2] represents the *control strategy*.

Two cases are possible: the *levers* can be *mutually dependent* or *independent*:

(a) $X1_t$ and $X2_t$ are *mutually dependent* if the manager cannot regulate the former independently of the latter because the two levers vary in the "opposite" direction with respect to the variations in Y_t. Thus, a variation in $X1_t$ in sense "**s**" normally implies as well a variation in $X2_t$ in the opposite *direction* "**o**," even if the variations are different in size; this type of control is defined as a *constrained-variable* (or opposed) control.

(b) $X1_t$ and $X2_t$ are *independent* when management can decide whether or not to modify only the first, second, or both levers; there is thus a *free variable* (or concurrent) control.

The simplest example of a *mutually dependent dual-lever* control system is that which regulates the water level (Y_t) in a tub or a tank by the use of two opposite control "levers," as illustrated in Fig. 4.1. Lever $X1_t$ indicates the width of the drain opening and the amount of time it remains open; lever $X2_t$ is the rotation of the water faucet that regulates the water flow input. If the manager is rational then it is intuitively clear that when there is an inflow of water from an opening of the faucet there cannot at the same time be an outflow through the drain (unless there is some distraction). Similarly, when the water drains out, it is useless to open the water inflow faucet.

Thus, in this control system the manager must manipulate the faucet and drain in the *opposite direction of variation*, until one of the two is closed. To indicate this constraint, I have inserted in the CLD in Fig. 4.1 the *pair of dashed arrows*, in direction "**o**," representing the reinforcing *loop* indicated by [**R**], whose effect is to indicate to the manager that he or she must rotate the two control levers in the *opposite* direction.

The water level dynamics in the bathtub, with correlated inflows and outflows and discrete-variation levers, following the strategy shown in Fig. 4.1 of manipulating the input and output levers in the opposite direction, is illustrated in Fig. 4.2. The control panel at the top of Fig. 4.2 shows the parameters that indicate the construction specifications chosen by the designer (the speed and section of X1 and X2).

I define the output lever X1 which causes a decrease in Y_t as a *negative* lever; each notch of the output lever means that X1 reduces the water level with a negative outflow of $g_1 = 5$ units of water (maximum); this value derives from the product of two parameters: the *speed*, $vg_1 = 1$, that is, the intensity of the liquid flowing out for each notch in the open direction, and the *section*, $sg_1 = 5$, which instead
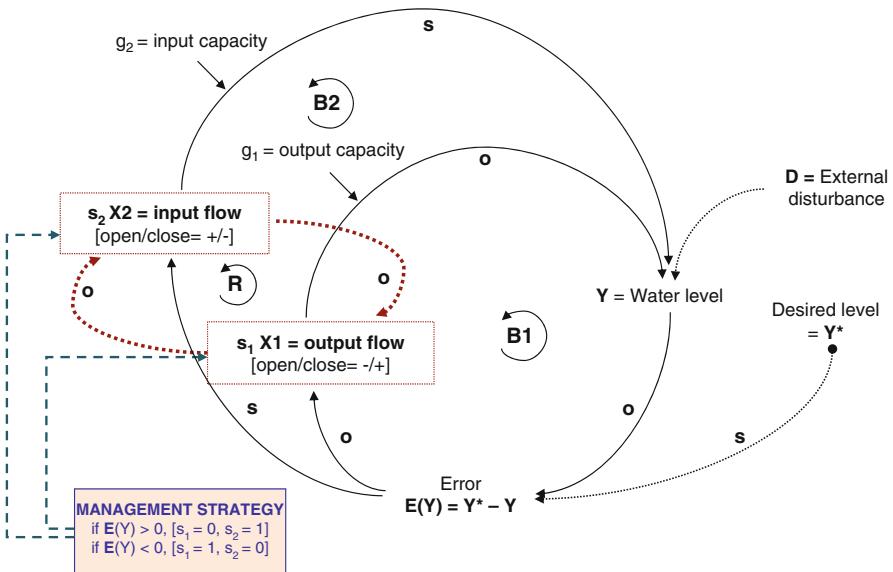
Fig. 4.1 Control of the water level with two *mutually dependent* levers

indicates how many notches on the output lever can be kept open at any instant, thus expressing the maximum *section* of the drain hole. The product of the *section* and the *velocity*, $g_1 = 5$, quantifies the maximum *capacity* flow of the lever output, that is, the maximum amount of the outflow that reduces the water level at each instant.

The parameters for the input lever X2—which produces the inflow—have a similar meaning. Since it represents a *positive* value, each notch on the input lever increases the water level by $g_2 = 4$ units, since X2 has a water *speed* of $vg_2 = 2$; the lever can be opened as far as the lever *section* $sg_2 = 2$ at each instant. The maximum *capacity* flow of the lever input is thus $g_2 = 4$ units per instant.

Since in Fig. 4.2 I have assumed the water level must be reduced from $Y_0 = 100$ to $Y^* = 40$ units (initial objective), the system manager acts rationally if, at $t = 1$, he or she opens the lever $X1_1$ to its maximum capacity, $g_1 = 5$ units (maximum admissible section and maximum admissible velocity); as a consequence, at $t = 1$ the level falls from $Y_0 = 100$ to $Y_1 = 95$ units of water.

The manager could have opened lever $X1_1$ by only one notch or chosen a lower velocity; but this would not have been rational, since it would have slowed down the achievement of the objective.

The level $Y^* = 40$ units is reached at $t = 12$. At $t = 14$ the governance changes the objective, which makes it necessary to increase the level from $Y^* = 40$ to $Y^* = 80$ units. Thus, the inflow lever must be used, which has a *water flow capacity* of $g_2 = 4$ units, thereby filling the tub at instant $t = 24$.

To complete this simple simulation, we assume that at $t = 26$ the governance decides to decrease the water level from $Y^* = 80$ to $Y^* = 30$ units; this new objective is achieved at $t = 36$.

Levers					
output lever X1		input lever X2			
section	5	section	2	$Y_0 =$	100
speed	1	speed	2	$Y^* =$	40
capacity	5	capacity	4		

t	OUTFLOW - X1	INFLOW - X2	LEVEL = Y	Y* = OBJECTIVE	Δ OBJECTIVE	DISTURBANCES	ERROR = Y* - Y
0	0.0	0.0	100.0	40.0	0.0	0.0	-60.0
1	5.0	0.0	95.0	40.0	0.0	0.0	-55.0
2	5.0	0.0	90.0	40.0	0.0	0.0	-50.0
3	5.0	0.0	85.0	40.0	0.0	0.0	-45.0
4	5.0	0.0	80.0	40.0	0.0	0.0	-40.0
5	5.0	0.0	75.0	40.0	0.0	0.0	-35.0
6	5.0	0.0	70.0	40.0	0.0	0.0	-30.0
7	5.0	0.0	65.0	40.0	0.0	0.0	-25.0
8	5.0	0.0	60.0	40.0	0.0	0.0	-20.0
9	5.0	0.0	55.0	40.0	0.0	0.0	-15.0
10	5.0	0.0	50.0	40.0	0.0	0.0	-10.0
11	5.0	0.0	45.0	40.0	0.0	0.0	-5.0
12	5.0	0.0	40.0	40.0	0.0	0.0	0.0
13	0.0	0.0	40.0	40.0	0.0	0.0	0.0
14	0.0	0.0	40.0	80.0	40.0	0.0	40.0
15	0.0	4.0	44.0	80.0	0.0	0.0	36.0
16	0.0	4.0	48.0	80.0	0.0	0.0	32.0
17	0.0	4.0	52.0	80.0	0.0	0.0	28.0
18	0.0	4.0	56.0	80.0	0.0	0.0	24.0
19	0.0	4.0	60.0	80.0	0.0	0.0	20.0
20	0.0	4.0	64.0	80.0	0.0	0.0	16.0
21	0.0	4.0	68.0	80.0	0.0	0.0	12.0
22	0.0	4.0	72.0	80.0	0.0	0.0	8.0
23	0.0	4.0	76.0	80.0	0.0	0.0	4.0
24	0.0	4.0	80.0	80.0	0.0	0.0	0.0
25	0.0	0.0	80.0	80.0	0.0	0.0	0.0

Bathtub controlled by two mutually dependent faucets **Dynamics of the input and output control levers**

Fig. 4.2 Dynamics of a bathtub controlled by two *mutually dependent* faucets

The graph of the water level dynamics can easily be produced, as shown in Fig. 4.2; the graph reveals the best *strategy* (see also Fig. 4.1) for varying the water level in the tub: opening only the output lever for water *outflow* or the input lever for water *inflow*. Nevertheless, the manager could have made other, illogical choices; for example, he or she could have tried to empty or fill the tub by simultaneously opening, for several periods, both the outflow and inflow levers. To avoid the risk of overflowing, designers usually give the outflow lever a greater overall capacity than the inflow one, as in the example in Fig. 4.2. To an outside observer such choices would appear irrational, because they are inefficient. Thus, the dynamics illustrated in Fig. 4.2 can be considered to be produced by an *automatic* control system that adopts the best *strategy* for achieving its objectives.

The simple control system for a tub with interdependent levers can be generalized in a number of ways, since it can be compared to a vast number of systems that entail a level that varies on the basis of inputs and outputs over different periods of time: warehouses, ships, stocks, etc. In Sect. 4.3 we will use this control system to simulate the behavior of other systems.

Technical note. The control system in Fig. 4.2 is a *dual-lever system*, but if it had been designed to allow the manager to control both the *speed* and *section* of the two levers, then the system would have become a *four-lever system* composed of two *mutually dependent* pairs of levers: one pair that regulates the input flows and the other the output flows. However, within each pair, the levers of speed and section would be *independent* of each other, and the definition of the strategy to adopt would be less simple.

4.2   Dual-Lever Control System with Independent Levers: Control Strategy

The simplest example of *independent* "dual-lever" control system—as defined in point (b) of previous section—is the regulation of the shower temperature using two separate faucets: one, $(X1_t = H_t)$, to regulate the flow of hot water, and the other, $(X2_t = C_t)$, to regulate the cold water flow. The water temperature $(Y_t = W_t)$ derives from a *mix* of H_t and C_t. If the water is too cold the manager opens H_t or closes C_t; if it becomes too hot he or she can close H_t or open C_t or adjust both faucets simultaneously until he or she obtains the desired temperature H^*.

Figure 4.3 represents the *logical structure* of a shower with two separate faucets to control water temperature, which is a general version of the model in Fig. 2.7; the hot and cold sensations felt by the person taking the shower generate the error temperature, which I have indicated as $E(W) = W^* - W_t$. The *governor-manager* of the system (the person taking the shower) can decide, after calculating and evaluating the temperature error (using his or her own skin), how to vary the rotation of the faucets, taking into account the temperature deviation rates, g_1 and g_2 (assumed to be constant), produced by each degree of rotation of the faucets.
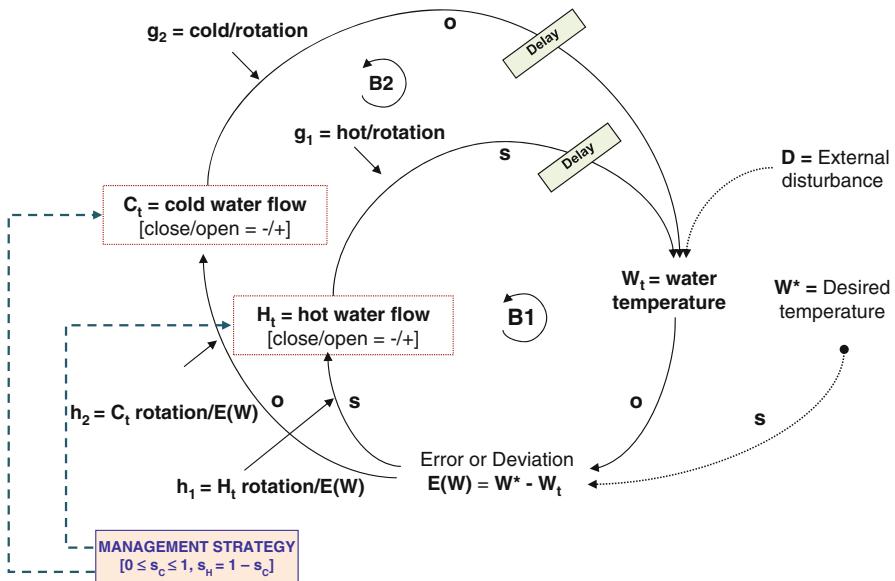
Fig. 4.3 Dual-lever temperature control

Let us assume (though each of us may choose differently) that the *governor-manager*, upon feeling the water too hot, prefers to play it safe and *always* close the hot water faucet first, only subsequently opening the cold water one, if the first decision does not solve the problem; when the water is too cold, the manager chooses to open the hot water faucet and then regulate the cold water one. Thus, the manager has chosen a precise *control strategy* that can be expressed as follows: "if you feel a significant difference between the water temperature detected by your skin (W) and the optimal temperature (W^*), first regulate faucet H_t and only subsequently faucet C_t." How many of us would share this strategy? How many would prefer to act simultaneously on the two faucets so as to speed up the regulation process? Conclusion: The control with two *free variables always* implies a *strategy* that defines an *order of priorities* of action on the *control levers* or the choice of a *mix* of the two levers based on a specific vector $[s_1, s_2]$ that operationalizes the *control strategy*; these options do not depend on the structure of the control system but on the information available to the manager and on his or her experience.

The term *strategy* used in the context of control refers to "a pattern of regulation of the control levers established by the manager based on his experience," after having determined that what was successful in the past can lead to success in the future. This definition is perfectly in line with the classical term of strategy as "*the employment of battles to gain the end of war*" coined by the Prussian general, Carl von Clausewitz. According to this view, *strategy* is the set of actions (control levers) by which the army commander (manager), based on his knowledge and

experience, uses the available arms and his soldier's skills (effector) to defeat the enemy in battle (control lever) in order to defend or conquer a territory for the well-being of his country (objectives).

Today the term is used in a more general sense in all fields where it is necessary to decide how to activate a group of action variables (Xn) in order to achieve a set of objectives (Y*) given the presence of adversaries or competitors or even other outside disturbances. It is used in this sense in *game theory* to indicate the *rule* or the *frequency* with which the various courses of action must be chosen during the successive repetitions of a game in order to maximize the winnings.

In management theory, Johnson and Scholes define strategy, as follows:

> Strategy is the direction and scope of an organisation over the long-term: which achieves advantage for the organisation through its configuration of resources within a challenging environment, to meet the needs of markets and to fulfill stakeholder expectations (Johnson and Scholes 2002, p. 3).

Chandler also includes the determination of the objectives in the definition of strategy:

> Strategy can be defined as the determination of the basic long-term goals and objectives of an enterprise, and the adoption of courses of action and the allocation of resources necessary for carrying out these goals (Chandler 1962, p. 13).

The term is well suited to multi-lever control systems, since there is a group of control levers that must be manipulated in a sequence of "moves and counter-moves" by assigning positive or negative values, with a certain frequency and with successive repetitions of the control process, so as to achieve the objective with the maximum efficiency, in the minimum amount of time, and with the minimum use of resources.

While in *manual* control systems the strategy must be decided each time and changed during the period of control, the creation of an *automatic* control system with *two independent levers* requires by nature that the *control strategy* adopted be defined *in advance*. An example of an *automatic* control system of water temperature is shown in the Excel simulation in Fig. 4.4, which presents a segment of the numerical dynamics—for 35 instants—of the water temperature variable, W_t, for a perfectly *symmetrical* shower, assuming two delays in adjusting the temperature and a reaction time of 1. The dynamics are determined by the choices indicated in the initial control panel. As the control panel shows, I have assumed that each notch of the hot water faucet increases (if opened) or reduces (if closed) the water temperature by $g_1 = 1°$; likewise, the cold water faucet reduces (if opened) or increases (if closed) the temperature by $g_2 = 1°$. The control *strategy* was determined in this example as follows: since the two faucets can be simultaneously regulated, if the *governor-manager* wants to rapidly reduce the temperature from $W_0 = 35\,°C$ (initial temperature) to $W^* = 10\,°C$ (objective), he or she uses both the hot water "lever" (second column), which is closed 2.5 notches (with each notch having a range of $10°$), and the cold water one, which instead is open 2.5 notches. A similar strategy is used if he or she needs to increase the temperature.

H hot lever	1	C cold lever	-1	$W_0 =$	35
hot treshold	5	cold treshold	3	$W^* =$	10
r (X)	1	capacity	10		

t	H = HOT LEVER	C = COLD LEVER	WATER TEMPERATURE	W^* = OBJECTIVE	Δ OBJECTIVE	DISTURBANCES	ERROR = W^* - W
0	0.0	0.0	35.0	10.0	0.0	0.0	-25.0
1	-2.5	2.5	35.0	10.0	0.0	0.0	-25.0
2	-2.5	2.5	35.0	10.0	0.0	0.0	-25.0
3	-2.5	2.5	30.0	10.0	0.0	0.0	-20.0
4	-2.0	2.0	25.0	10.0	0.0	0.0	-15.0
5	-1.5	1.5	20.0	10.0	0.0	0.0	-10.0
6	-1.0	1.0	16.0	10.0	0.0	0.0	-6.0
7	-0.6	0.6	13.0	10.0	0.0	0.0	-3.0
8	-0.3	0.0	11.0	10.0	0.0	0.0	-1.0
9	-0.1	0.0	9.8	10.0	0.0	0.0	0.2
10	0.0	0.0	9.5	10.0	0.0	0.0	0.5
11	0.0	-0.1	9.4	10.0	0.0	0.0	0.6
12	0.0	-0.1	9.4	10.0	0.0	0.0	0.6
13	0.0	-0.1	9.5	10.0	0.0	0.0	0.5
14	0.0	-0.1	9.5	10.0	0.0	0.0	0.5
15	0.0	0.0	9.6	10.0	0.0	0.0	0.4
16	0.0	0.0	9.6	10.0	0.0	0.0	0.4
17	0.0	0.0	9.7	10.0	0.0	0.0	0.3
18	0.0	0.0	9.7	30.0	20.0	0.0	20.3
19	2.0	-2.0	9.8	30.0	0.0	0.0	20.2
20	2.0	-2.0	9.8	30.0	0.0	0.0	20.2
21	2.0	-2.0	13.9	30.0	0.0	0.0	16.1
22	1.6	-1.6	17.9	30.0	0.0	0.0	12.1
23	1.2	-1.2	21.9	30.0	0.0	0.0	8.1
24	0.8	-0.8	25.2	30.0	0.0	0.0	4.8
25	0.0	-0.5	27.6	30.0	0.0	0.0	2.4
26	0.0	-0.2	29.2	30.0	0.0	0.0	0.8
27	0.0	-0.1	29.7	30.0	0.0	0.0	0.3
28	0.0	0.0	29.9	30.0	0.0	0.0	0.1
29	0.0	0.0	30.0	30.0	0.0	0.0	0.0
30	0.0	0.0	30.0	30.0	0.0	0.0	0.0
31	0.0	0.0	30.0	20.0	-10.0	0.0	-10.0
32	-1.0	1.0	30.0	20.0	0.0	0.0	-10.0
33	-1.0	1.0	30.0	20.0	0.0	0.0	-10.0
34	-1.0	1.0	28.0	20.0	0.0	0.0	-8.0
35	-0.8	0.8	26.0	20.0	0.0	0.0	-6.0

Fig. 4.4 Shower controlled by two independent and continuous faucets with two delays

Nevertheless, we can observe that the use of the two levers is not perfectly symmetrical since, in this specific example, the user-manager prefers to regulate both levers to rapidly lower the temperature when it is too high and must be lowered; but when the temperature, at $t = 7$, reaches the "cold threshold" of
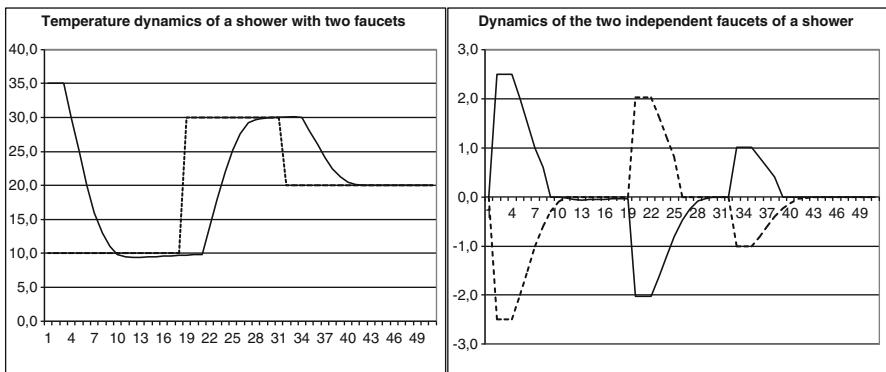
Fig. 4.5 Dynamics of the variables of the control system for the shower in Fig. 4.4

$E(W)_7 = 3°$, he or she regulates the temperature by only acting on the hot water lever (closing it slowly) to avoid the risk of lowering the temperature below the objective.

Vice versa, when the temperature must be raised (which occurs in period $t = 18$), the manager manipulates both levers so that the water quickly heats up; but when the "hot threshold" of $E(W)_{24} \approx 5°$ is reached ($t = 24$) he or she prefers to reach the objective by using the cold water lever (closing it slowly) so as not to risk exceeding the objective. Of course this is not the only possible strategy, but it reflects our natural cautionary tendency, in this case not to risk "exceeding" the tolerance limits of hot and cold.

The left-hand table in Fig. 4.5 indicates the dynamics of the water temperature. The right-hand table indicates the trends in the opening and closing of the faucets, which represent the two control levers.

I think it is important to emphasize that in many cases dual-lever control systems are "*badly designed*," since they are programmed to always give preference to a single control lever, which causes errors in regulation; only intervention by an outside manager with considerable experience and knowledge can avoid the error that leads to irreversible damage to the behavior of the entire system.

We can become aware of this problem by considering, for example, a biological control system we constantly have to deal with, often in a rather problematic way: the control system of *thirst*, which can be represented by the model in Fig. 4.6. Among the many biological constraints in our body is that of the optimal or the normal *density standard* of salt, S, with respect to the volume of water, W, which is indicated in Fig. 4.6 by $D^* = (S/W)^*$; this represents the *objective* (or constraint) the system must achieve and maintain.

When the effective salt density (D_t) does not coincide with the standard one (D^*), the error $E(D) = D^* − D_t$ manifests itself in our organism as a *sensation of thirst*. The higher the difference ($D^* − D_t$), the more intense is the *thirst stimulus*. Thus, Fig. 4.6 represents a *dual-lever* control system. The ratio D_t between salt and
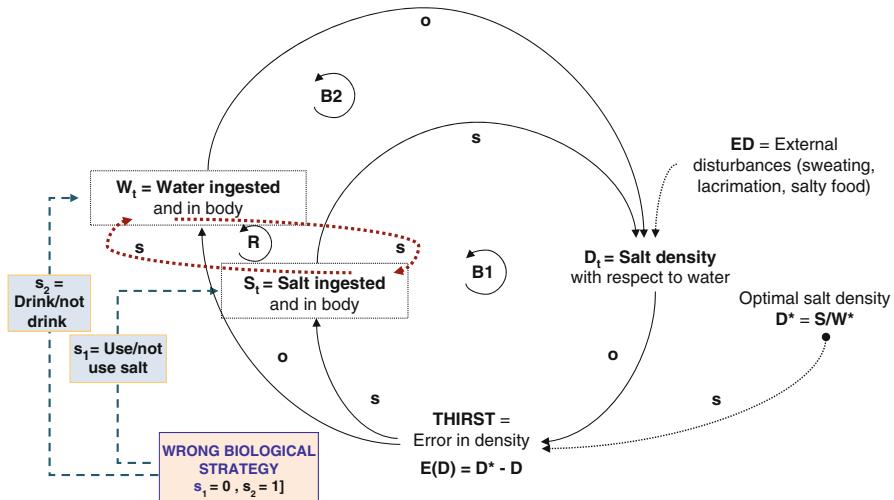
Fig. 4.6 Dual-lever salt water density control

water can be controlled by regulating the lever S_t for the amount of salt ingested, thereby forming the *loop* [**B1**], or the lever for the quantity of water ingested, W_t, or drinking lever, thereby forming the *loop* [**B2**].

In many cases the control system fails because it is "badly designed." In fact, the system cannot determine whether the *thirst* is caused by a scarcity/excess of water or by a scarcity/excess of salt. Thirst derives from $E(D)$, whether or not it is caused by the values of W or S.

Thus, thinking of our own experience, when we are thirsty we always use the single control lever W: "drink water." If the error, $E(D)$, is due to an excess of salt in the numerator (too much salt and too little water, because we have eaten salty foods), then by acting on the W_t lever we can reduce and eliminate thirst (we must take into account delays). However, if the sensation of burning thirst is from a lack of salt (we have sweated or cried too much, causing the salt to be expelled from our body along with the liquids), then the signal from the error $E(D)$ (too little salt and too much water) leads us, unfortunately, to drink again in order to try and eliminate our thirst. The intake of liquids further worsens D_t and increases the error, causing greater thirst and desire to consume liquids. This goes on until the system breaks down with … water indigestion. This is well known by people who live in warm-weather countries or who work in the sun; these managers know this simple strategy: "if, even when drinking, our burning thirst becomes unbearable, it is useless to insist on drinking; it is better to take a packet of minerals salts."

In the animal kingdom this system appears to be well designed in some cases. Nocturnal porcupines, that well-known rodent with long, pointy spines, search out salty foods, licking and eating anything covered with salt, from household goods to tires that have gone over salt spread on the road! Elephants consume grasses, small plants, bushes, fruit, twigs, tree bark, and roots. To supplement the diet, elephants

will dig up earth to obtain salt and minerals. Over time, African elephants have hollowed out deep caverns in a volcano mountainside on the Ugandan border to obtain salt licks and minerals (http://www.seaworld.org/animal-info/info-books/ elephants/diet.htm). Finally, we can note the periodic migration of zebras in Africa: each year they cover hundreds of kilometers to reach the Botswana's salt pans in order to eat salt and stock up for the rest of the year.

Another control system that, though badly designed, operates in a perfectly logical manner is the one that regulates our body temperature when we have a fever. An increase in temperature should produce significant sweating; however, as we have all observed and experienced, a fever makes us feel cold and gives us chills, which, in serious cases in young children, even leads to feverish convulsions.

It appears that the control system for body temperature is not well designed. In fact, however, from a logical point of view it functions perfectly. We know that chills are one of the physiological levers to combat outside cold with respect to the normal body temperature. Chills serve to produce heat in the muscles when our sensors perceive too low an environmental temperature. If we reflect on it, even when we are feverish our temperature sensors perceive an environmental temperature below our body temperature, which has risen due to the fever; as a result the body chill lever is triggered along with the one which causes us to look for blankets to reduce the sensation of cold. Thus, when we have a fever the natural control system for body temperature appears to be badly designed, since it cannot distinguish whether or not the difference $E(Y)$ between the body temperature, Y^*, and the environmental temperature, Y, is due to the fact that the latter is too low with respect to our equilibrium temperature or that our body temperature is too high with respect to the environmental one. In order to overcome the poor functioning of this natural control system the outside manager (doctor, family member, etc.) must intervene to activate the auxiliary levers to rectify the natural ones, for example, by placing cold water packs on the fevered body, taking off clothing, or taking medicine to lower the fever. Obviously, when the fever falls rapidly the natural control system acts in the opposite way to produce sweating and the sensation of heat, which causes the fevered individual to throw off his blankets.

4.3   Impulse Control Systems

The water tanks and shower operate with *two levers*, X1 and X2, that control Y through gradual variations over time; even if measured at discrete instants, the levers usually take on definite values for each $t = 0, 1, 2, \ldots$, which are included in the observation interval T. Such systems can be defined as *gradual* or *continuous functioning* control systems, depending on whether the control levers admit a discrete or a continuous functioning.

Let me now introduce another important class of dual-lever control systems: those where one control lever produces a sequence of inputs or outputs for Y_t over a certain period, while the other produces a single value, opposite in direction of

variation and resembling an *impulse*, at the beginning or the end of the period. More precisely, a system operates by *impulses* when a control lever—for example, X1—has a nonzero value as soon as Y_t, or the variance $E(Y)_t$, reaches a given level (previously defined as the *objective*) while maintaining the value $X1_t = 0$ for every other level of Y_t or E_t. The second control lever, X2, instead takes on the value zero if the other variable has a nonzero value and has a value opposite to that of the objective when X1 becomes zero. This joint behavior produces a periodic variation of Y from zero to the value of the objective. To show this, let us take as an example an ideal warehouse where stocks of whatever kind are accumulated; these stocks increase by means of new inputs and decrease as a result of the output.

We let Y_t indicate the variable that expresses the stock's trends, X1 the *output* variable, and X2 the *input* variable. We assume that the two levers are interdependent, as in the case of the bathtub. Impulses can function in two ways: the initial stock is at a level equal to Y_{min} and is increased gradually by the X2 lever until it reaches the objective Y^*; the stock, through the action of the X1 lever, then takes on the level $Y_t = 0$, from which it begins to gradually increase until it once again reaches the objective Y^*. The cycle repeats itself until period T is over. We can call this ideal *impulse* control system, CI-PO, an acronym for *continuous input–point output*, which immediately calls to mind its operational method.

Figure 4.7 shows numerically the first two cycles of a control system that functions as I have just described. I have not included variations in objectives or disturbances, but these would only produce a temporary alteration of the cycles, which would regularly repeat themselves once the disturbance has ended. We observe from the control panel that the objective is $Y^* = 100$ and that the input variables X2, with a *velocity* of $v_2 = 2$ and a *section* of $s_2 = 10$, resupply the warehouses with shipments of *capacity* $g_2 = 20$ units, which produces the same increase in the stock for each t_h. When Y_t reaches $Y^* = 100$ units it is "discharged" with a single output produced by the X1 lever, whose capacity is $g_1 = 100$ units; the stock is eliminated and starts to increase anew.

Figure 4.8 illustrates numerically the consequences of raising the objective, which occurs at $t = 9$, as well as of decreasing disturbances in the instants from $t = 18$ to $t = 22$. Despite these disturbances the control system always succeeds in achieving the value $Y^* = 100$.

We can imagine an alternative form of functioning of the ideal warehouse viewed as a control system: the warehouse starts at a certain level and is gradually decreased through the X1 lever until it reaches the value $Y_t = 0$, after which it is increased by the X2 lever until it reaches the restocking value $Y_0 = 100$. After the increase the stock then gradually diminishes through the action of the X1 lever. I propose naming this type of ideal control system a PI–CO system, which stands for *point input–continuous output*, in order to succinctly highlight its operational logic.

We can improve the model of the PI–CO system by adding the constraint that the resupply occurs not at stock level $Y_t = 0$ but at a positive level, $Y_{min} > 0$, which I call the *security level*; this takes on the significance of the objective Y^* that the system must achieve and maintain as a minimum level. Figure 4.9 shows

OUT	X1 velocity	10	X1 section	10	initial stock	Y_0 =	0
IN	X2 velocity	2	X2 section	10	point output	Y^* =	100

| t	OUTPUT LEVER X1	INPUT LEVER X2	STOCK Y_T	OBJECTIVE Y^*	Δ OBJECTIVE	D = DISTURBANCES	ERROR E(Y)$_t$ = Y^* - Y^t
0	0.0	0.0	0.0	100.0	0.0	0.0	100.0
1	0.0	20.0	20.0	100.0	0.0	0.0	80.0
2	0.0	20.0	40.0	100.0	0.0	0.0	60.0
3	0.0	20.0	60.0	100.0	0.0	0.0	40.0
4	0.0	20.0	80.0	100.0	0.0	0.0	20.0
5	0.0	20.0	100.0	100.0	0.0	0.0	0.0
6	-100.0	0.0	0.0	100.0	0.0	0.0	100.0
7	0.0	20.0	20.0	100.0	0.0	0.0	80.0
8	0.0	20.0	40.0	100.0	0.0	0.0	60.0
9	0.0	20.0	60.0	100.0	0.0	0.0	40.0
10	0.0	20.0	80.0	100.0	0.0	0.0	20.0
11	0.0	20.0	100.0	100.0	0.0	0.0	0.0
12	-100.0	0.0	0.0	100.0	0.0	0.0	100.0
13	0.0	20.0	20.0	100.0	0.0	0.0	80.0
14	0.0	20.0	40.0	100.0	0.0	0.0	60.0
15	0.0	20.0	60.0	100.0	0.0	0.0	40.0
16	0.0	20.0	80.0	100.0	0.0	0.0	20.0
17	0.0	20.0	100.0	100.0	0.0	0.0	0.0
18	-100.0	0.0	0.0	100.0	0.0	0.0	100.0
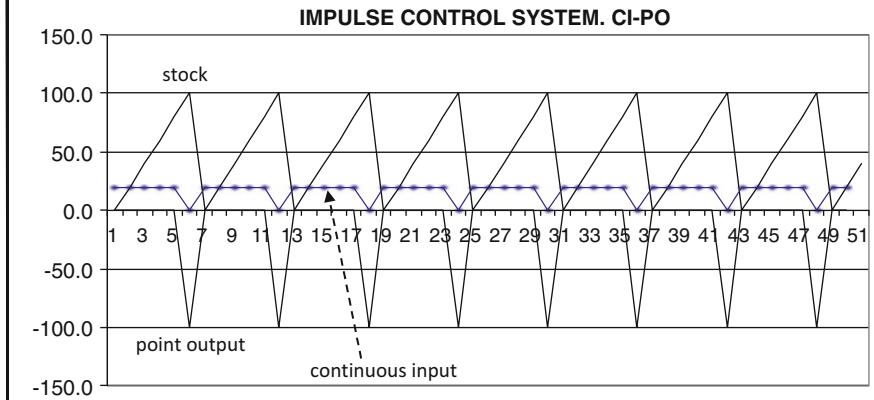
Fig. 4.7 Impulse control system with CI–PO functioning, discrete levers, and no disturbances

numerically the first two cycles of a warehouse that functions in this way; it also assumes a variation in the objective in period $t = 18$.

At first sight, control systems that operate *by impulses* may seem somewhat abstract; but even after quick reflection we can easily realize how common they are. A typical CI-PO control system is our toilet. When we push the button to flush the toilet the water drains out almost instantaneously (PO); after the tank is emptied it starts to fill again (CI), until the sensor and regulation apparatus, already shown in Fig. 3.13, stop the water level from rising after reaching the maximum level for the tank.

OUT	X1 velocity	10	X1 section	10	initial stock	Y_0 =	0
IN	X2 velocity	2	X2 section	10	point output	Y^* =	100

t	OUTPUT LEVER X1	INPUT LEVER X2	STOCK Y_T	OBJECTIVE Y^*	Δ OBJECTIVE	D = DISTURBANCES	ERROR $E(Y)_t$ = $Y^* - Y^t$
0	0.0	0.0	0.0	100.0	0.0	0.0	100.0
1	0.0	20.0	20.0	100.0	0.0	0.0	80.0
2	0.0	20.0	40.0	100.0	0.0	0.0	60.0
3	0.0	20.0	60.0	100.0	0.0	0.0	40.0
4	0.0	20.0	80.0	100.0	0.0	0.0	20.0
5	0.0	20.0	100.0	100.0	0.0	0.0	0.0
6	-100.0	0.0	0.0	100.0	0.0	0.0	100.0
7	0.0	20.0	20.0	100.0	0.0	0.0	80.0
8	0.0	20.0	40.0	100.0	0.0	0.0	60.0
9	0.0	20.0	60.0	150.0	50.0	0.0	90.0
10	0.0	20.0	80.0	150.0	0.0	0.0	70.0
11	0.0	20.0	100.0	150.0	0.0	0.0	50.0
12	0.0	20.0	120.0	150.0	0.0	0.0	30.0
13	0.0	20.0	140.0	150.0	0.0	0.0	10.0
14	0.0	10.0	150.0	150.0	0.0	0.0	0.0
15	-150.0	0.0	0.0	150.0	0.0	0.0	150.0
16	0.0	20.0	20.0	150.0	0.0	0.0	130.0
17	0.0	20.0	40.0	150.0	0.0	0.0	110.0
18	0.0	20.0	50.0	150.0	0.0	-10.0	100.0
19	0.0	20.0	50.0	150.0	0.0	-20.0	100.0
20	0.0	20.0	40.0	150.0	0.0	-30.0	110.0
21	0.0	20.0	40.0	150.0	0.0	-20.0	110.0
22	0.0	20.0	50.0	150.0	0.0	-10.0	100.0
23	0.0	20.0	70.0	150.0	0.0	0.0	80.0
24	0.0	20.0	90.0	150.0	0.0	0.0	60.0
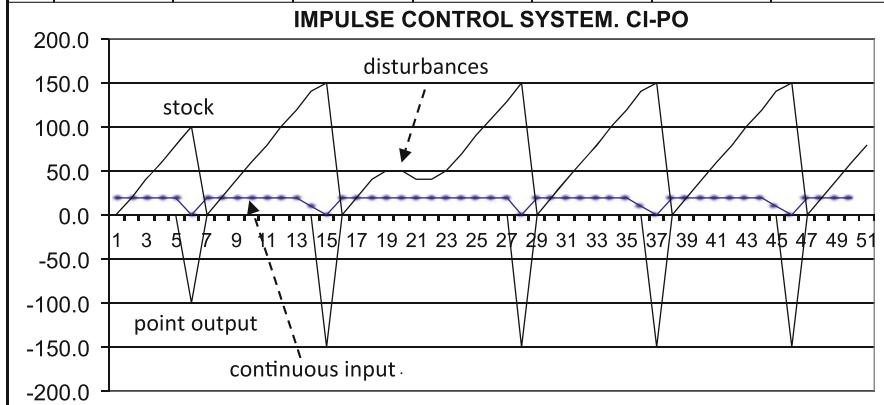
Fig. 4.8 Control system with CI–PO functioning, with disturbances

The velocity with which the tank fills up depends, as always, on the *capacity* of lever X1, which in turn depends on the tube *section* and the water *velocity*. We can also view plans for capital accumulation as CI-PO control systems. Most CI-PO control systems are observed in the living world; fruit slowly grows (CI) with the

| OUT | X1 velocity | 3 | X1 section | 5 | initial stock | $Y_0 =$ | 100 |
| IN | X2 velocity | 10 | X2 section | 10 | security level | $Y^* =$ | 20 |

t	OUTPUT LEVER X1	INPUT LEVER X2	STOCK Y_T	OBJECTIVE Y^*	Δ OBJECTIVE	D = DISTURBANCES	ERROR E(Y)$_t$ = Y^* - Y^t
0	0.0	0.0	100.0	20.0	0.0	0.0	-80.0
1	-15.0	0.0	85.0	20.0	0.0	0.0	-65.0
2	-15.0	0.0	70.0	20.0	0.0	0.0	-50.0
3	-15.0	0.0	55.0	20.0	0.0	0.0	-35.0
4	-15.0	0.0	40.0	20.0	0.0	0.0	-20.0
5	-15.0	0.0	25.0	20.0	0.0	0.0	-5.0
6	-5.0	0.0	20.0	20.0	0.0	0.0	0.0
7	0.0	100.0	120.0	20.0	0.0	0.0	-100.0
8	-15.0	0.0	105.0	20.0	0.0	0.0	-85.0
9	-15.0	0.0	90.0	20.0	0.0	0.0	-70.0
10	-15.0	0.0	75.0	20.0	0.0	0.0	-55.0
11	-15.0	0.0	60.0	20.0	0.0	0.0	-40.0
12	-15.0	0.0	45.0	20.0	0.0	0.0	-25.0
13	-15.0	0.0	30.0	20.0	0.0	0.0	-10.0
14	-10.0	0.0	20.0	20.0	0.0	0.0	0.0
15	0.0	100.0	120.0	20.0	0.0	0.0	-100.0
16	-15.0	0.0	105.0	20.0	0.0	0.0	-85.0
17	-15.0	0.0	90.0	20.0	0.0	0.0	-70.0
18	-15.0	0.0	75.0	40.0	20.0	0.0	-35.0
19	-15.0	0.0	60.0	40.0	0.0	0.0	-20.0
20	-15.0	0.0	45.0	40.0	0.0	0.0	-5.0
21	-5.0	0.0	40.0	40.0	0.0	0.0	0.0
22	0.0	100.0	140.0	40.0	0.0	0.0	-100.0
23	-15.0	0.0	125.0	40.0	0.0	0.0	-85.0
24	-15.0	0.0	110.0	40.0	0.0	0.0	-70.0

Fig. 4.9   Control system with PI–CO functioning, without disturbances

gradual help of heat and nutrients (variable X2); it then matures (Y^*), falls from the tree, or ends up on some table (PO). But the control system in the next seasonal cycle will once again transform the input of heat and nutrients into a new fruit.

The fact that, as a rule, pregnancy in humans lasts about 40 weeks is the clearest demonstration of the action of such systems, even if in this case the cycles are not automatic or regular.

PI–CO control systems are equally as common. Our pantry, fridge, and the gas tank in our car are under our daily observation. Unless we constantly visit hypermarkets or gas stations, we all supply ourselves with a stock of food, home products, and gas (PI) to last for several days; we use these goods regularly (CO), resupplying ourselves when their levels fall "dangerously" close to the security level, Y^*.

A good meal supplies us with the energy (PI) for several hours of work, during which it is used up (CO), thus making us desire a new meal (PI). This cycle repeats itself throughout our lifetime.

Systems that function not by impulses but gradually can, by extension and when appropriately construed, be viewed as *continuous input–continuous output* (CI–CO) systems.

4.4   Multi-lever Control Systems

The considerations in the preceding *sections* for *dual-lever* control systems are also valid for *multi-lever* systems with any number of levers, all or in part *mutually dependent* or *independent*. Obviously, the greater the number of levers the more the rational manager must determine an effective *strategy* that takes into account the urgency of achieving the objective and the *capacity* of each lever, in addition to the cost of regulation (for these aspects, see Sect. 4.7). Control without a strategy means "blindly navigating"; control is by "trial and error," and it is not easy to optimize the time required to achieve the objective. Often control systems are multi-lever because the levers have a limited range of action and, when the error is large, must be employed contemporaneously (Fig. 4.10) or in succession (Fig. 4.11) to achieve the objective. In these cases the strategy depends on the possibility of regulating the levers.

An example of this type of control is the *spontaneous control system* for focusing our eyes on objects in a horizontal plane with respect to the normal vertical position of the bust, based on the model in Fig. 4.10. This system is also suited to objects in any area that can be taken in by our gaze; however, if we wish to limit ourselves to the horizontal plane, the control system in Fig. 4.10 requires the action of only four control levers.

As we can all verify, when we have to focus on an object in a horizontal plane with about a 20° radius to the left or the right with respect to the center, we are "naturally" inclined to horizontally rotate our eyeballs (E_t); for objects further to the outside, within about a 70° radius, focusing also requires the rotation of the head (H_t). For objects even further to the outside, up to about a 90° radius left or right, we must rotate our bust (B_t). Beyond this limit, if we want to focus our gaze on an object "behind" us we are forced to turn our entire body (BO_t). Under normal conditions we all adopt this *natural strategy* of four-lever regulation: first rotating the eyes, then the head, then the bust, and finally, if necessary, our body, according to the percentages to apply to the reaction rates "*h*" to regulate the various levers, as
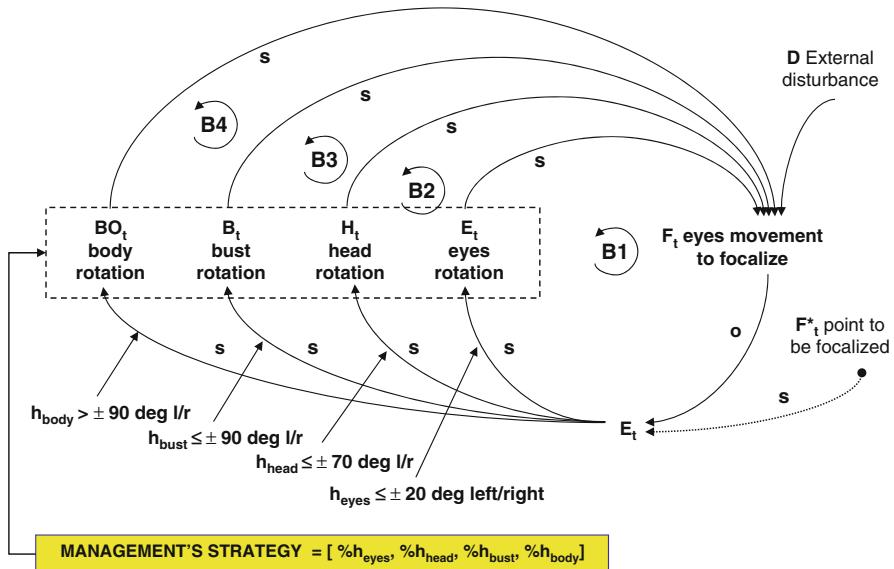
Fig. 4.10 Focusing on an object in a horizontal plane with a multi-lever control system
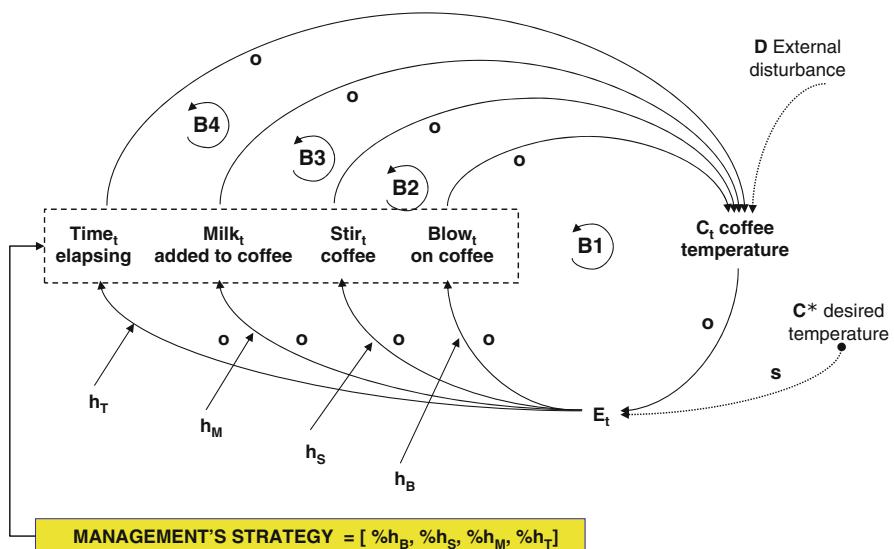
Fig. 4.11 Reducing the temperature of a liquid with a multi-lever control system

shown in the lower part of Fig. 4.10. In the model in Fig. 4.10 the levers are normally regulated in *cumulative fashion* by acting jointly in the order indicated by the strategy in relation to the amount of the error.

However, if the manager tries to regulate the system under "non-normal" conditions, he or she may be forced to find a different strategy. When I have a stiff neck I try to rotate my eyes as much as possible and then rotate my upper body, thereby avoiding rotating my head so as not to feel cervical pain. When I have a bad back I try not to activate the lever for rotating my bust, preferring instead to rotate my body by moving my feet. When I am driving my car I am quite limited in regulating the lever that rotates my bust, and the lever for rotating my body is completely deactivated. Fortunately the latter lever is usually substituted by the rotation of the eyes, which is possible thanks to the *supplementary lever* of the rear-view mirror, which produces a change in perspective and scale.

Another multi-lever system that we use daily is the control of the temperature of a liquid or of food that is too hot to be ingested, since its temperature exceeds that which our organism recognizes as the tolerance limit. Figure 4.11 represents a possible configuration of the control system for lowering the temperature of coffee served to us boiling hot.

When we bring the cup to our mouth and our sensors perceive the temperature of the coffee (C_t), which is too high compared to the desired temperature (C^*), we spontaneously activate an initial lever: blowing on the coffee to cool it (BLOW lever); a second, alternative lever is to stir the coffee several times with a cold spoon, thus exploiting the lower ambient temperature (STIR lever). If we are in a rush we can also add cold milk (MILK lever). On the contrary, if time allows we can simply wait for the necessary amount of time for the coffee to dissipate its incorporated heat to the atmosphere, thus lowering the temperature (TIME lever). These four levers can be used in succession, or one can be used as an alternative to another, or in a *mix* decided on by the manager-user. During a quick coffee break we would normally activate the BLOW and MILK levers. During a gala no one would dream of blowing on his or her coffee or stirring it repeatedly. The preferred lever could be to allow the time to go by.

We should note the difference between the control system in Fig. 4.10 and that in Fig. 4.11. The system in Fig. 4.10 is a typical *attainment* control system, whose levers can have positive and negative values. The system in Fig. 4.11 is a typical *halt* system; if the coffee becomes too cold it is undrinkable and the levers can be regulated, individually or jointly, in a single direction only (once the cold milk is added it cannot be removed).

A typical multi-lever control system which is particularly interesting because it is common in nature and has an infinite number of applications is that which controls the movement of a point toward an objective in an "*N*-dimensional space." For a space of $N = 4$ dimensions, defined by the coordinates $S = (x, y, z, w)$ placed over four theoretical axes, the system can be described as a control system to shift a point $P_t = (x_t, y_t, z_t, w_t)$, defined in S, toward a different point, assumed to be the objective, $P^* = (x^*, y^*, z^*, w^*)$, assumed to be fixed. In fact, the shifting of P toward P^* in the four-dimensional space entails the gradual elimination of the distance between P^* and P, according to the typical logic of control (Sect. 2.1).

In order to achieve the objective the control system must modify the position of the "intercepts" in each of the four theoretical axes.

A relatively simple, holonic-like process (Sect. 3.8) to produce the dynamics of the coordinates of P is to transform the single control system into four second-level control systems, considering each of the coordinates of the primary system objective, P^*, as a basic objective of the second-level systems. The control system for shifting P_t toward P^* is thus divided into four single-lever control systems, each of which is able to control the dynamics of a single coordinate of P_t, thereby determining and eliminating the error of the coordinate with respect to the value of the objective embodied in P^*. The result is the gradual shifting of P toward P^*, as shown in Fig. 4.12, which takes place over 101 periods with the parameters indicated in the control panel.

We must decide if and when P will reach P^* in order to allow the lower level systems to stop acting. For this we need to only go back to the first-level control system and calculate, for each instant, the Euclidean distance between P^* and P, representing the error $E(P)$, which is defined as

$$P^* - P = \sqrt{(x^* - x_t)^2 + (y^* - y_t)^2 + (w^* - w_t)^2 + (z^* - z_t)^2}$$

The *four-lever* first-level control system stops when $E(P) = P^* - P = 0$, which occurs at $t = 100$, when the four *single-lever*, second-level control systems, each acting independently, eliminate the distances in each coordinate (not shown in Fig. 4.12), thereby allowing the first-level system to shift the point from P to P^*.

In fact, Fig. 4.12 shows how the control system shifts the point of coordinates $P_0 = (0, -20, -50, -20)$ toward the *objective* point of coordinates $P^* = (180, 160, 150, 100)$ by means of the four second-level systems, each of which produces a shift along the axes, with the action rates "g" specified in the system's control panel; different action rates have been chosen for each pair of axes. The graph at the bottom of the table shows the dynamics of point P_t as it moves toward P^*, which is reached when the control system eliminates the distance between the two points. This movement is also shown by the curve representing the distance of P_t, calculated with *respect to the origin*. With these settings the four variables—x_t, y_t, z_t, w_t—are, in all respects, the four control levers the manager of the first-level control system regulates to produce the movement of P_t toward P^* by activating the second-level systems.

What strategy can the manager use to regulate the four levers (x_t, y_t, z_t, w_t) that control P? There are a number of possible strategies. Management could decide to modify the position of P by acting first on the "x" coordinate, then on "y," and subsequently on the other coordinates. It could also begin the movement of "w" and then act on "z," and so on. In the system in Fig. 4.12 it was decided to activate all four of the levers of P using the maximum value of "g" permitted for each (last line in the control panel). At time $t = 45$ the variables "x" and "y," which have higher *action rates* than the others ($g_x = 4$ and $g_y = 4$, as shown in the control panel), reach the objective values $x^* = 180$ and $y^* = 160$ (see the horizontal line at line $t = 45$).

		x	y	z	w	
	P	0	-20	-50	-20	initial position of point P
	P*	180	160	150	100	position of the objective point P*
	action rates - g	4	4	2	2	

t	g_X = traslat X	g_Y = traslat Y	g_Z = traslat Z	g_W = traslat W	x	y	z	w	x*	y*	z*	w*	dx=x*-x	dy=y*-y	dy=y*-y	dw=w*-w	distance of P from P*	Distance of P* from origin	Distance of P from origin
0					0	-20	-50	-20	180	160	150	100	180	180	200	120	345.25	300.83	57.45
1	4	4	2	2	4	-16	-48	-18	180	160	150	100	176	176	198	118	339.23	300.83	53.85
2	4	4	2	2	8	-12	-46	-16	180	160	150	100	172	172	196	116	333.23	300.83	50.79
3	4	4	2	2	12	-8	-44	-14	180	160	150	100	168	168	194	114	327.23	300.83	48.37
4	4	4	2	2	16	-4	-42	-12	180	160	150	100	164	164	192	112	321.25	300.83	46.69
5	4	4	2	2	20	0	-40	-10	180	160	150	100	160	160	190	110	315.28	300.83	45.83
6	4	4	2	2	24	4	-38	-8	180	160	150	100	156	156	188	108	309.32	300.83	45.83
[omitted]																			
44	4	4	2	2	176	156	38	68	180	160	150	100	4	4	112	32	116.62	300.83	247.75
45	4	4	2	2	180	160	40	70	180	160	150	100	0	0	110	30	114.02	300.83	253.97
46	0	0	2	2	180	160	42	72	180	160	150	100	0	0	108	28	111.57	300.83	254.85
47	0	0	2	2	180	160	44	74	180	160	150	100	0	0	106	26	109.14	300.83	255.76
48	0	0	2	2	180	160	46	76	180	160	150	100	0	0	104	24	106.73	300.83	256.69
49	0	0	2	2	180	160	48	78	180	160	150	100	0	0	102	22	104.35	300.83	257.66
50	0	0	2	2	180	160	50	80	180	160	150	100	0	0	100	20	101.98	300.83	258.65
[omitted]																			
60	0	0	2	2	180	160	70	100	180	160	150	100	0	0	80	0	80.00	300.83	270.00
61	0	0	2	0	180	160	72	100	180	160	150	100	0	0	78	0	78.00	300.83	270.53
62	0	0	2	0	180	160	74	100	180	160	150	100	0	0	76	0	76.00	300.83	271.06
[omitted]																			
96	0	0	2	0	180	160	142	100	180	160	150	100	0	0	8	0	8.00	300.83	296.92
97	0	0	2	0	180	160	144	100	180	160	150	100	0	0	6	0	6.00	300.83	297.89
98	0	0	2	0	180	160	146	100	180	160	150	100	0	0	4	0	4.00	300.83	298.86
99	0	0	2	0	180	160	148	100	180	160	150	100	0	0	2	0	2.00	300.83	299.84
100	0	0	2	0	180	160	150	100	180	160	150	100	0	0	0	0	0.00	300.83	300.83

Fig. 4.12 Dynamics of a point *P* that shifts toward an objective *P**

Thus in periods $t > 45$ their value is held at "0," since no further shift is necessary. The other two levers, with lower *action rates* ($g_z = 2$ and $g_w = 2$, as shown in the control panel), continue their trajectory until $t = 100$, when they reach their objective value, $w* = 150$ and $z* = 100$ (objective coordinates at *P**). Since the four second-level systems have attained their objectives, even the *four levers* stop acting. Precisely because of the differing dynamics of the four levers of the second-level

systems, at $t = 45$ the curve representing the distance of P from zero, together with that for the distance of P^* from P, shows a sudden variation.

If all the coordinates had infinitesimal and equal action rates, the optimal strategy would equate to the shifting of P_t toward P^* with a continuous straight movement along a straight line passing from P to P^*. The presence of "obstacles" between P and P^* and the use of different action rates might require the adoption of other strategies.

A concrete example of a control system that operates like the one in Fig. 4.12, in only a *three-dimensional* space, is the control of the position of the "movable point" of a crane that must move a load of bricks from point $P_0 = (x_0, y_0, z_0)$ to point $P^* = (x^*, y^*, z^*)$, for example, to the top of a roof. The moving point, $P_t = (x_t, y_t, z_t)$, shifts from P_0 to P^* by acting on the three coordinates through the angular rotation of the crane shaft (x_t), the horizontal shifting of the carriage (y_t), and the vertical shifting of the pulley (z_t). The control of these coordinates is carried out through the three single-lever, second-level control systems, which produce the variation in the P_t coordinates. Each of these is activated by a different effector (motor) and has a different *action rate*: the rotation of the crane is slow while that of the shifter of the horizontal and vertical axes is faster. The control *strategy* adopted by the crane operator to shift the load depends on the starting and ending position, but above all on his or her personal preferences and the constraints imposed by the "operations manual," which indicates the safety procedures to follow.

A new variant is the four-dimensional control system, which adds a fourth time dimension to the three spatial ones ($w = $ time). Any vehicle that travels in three-dimensional space in order to move from one point in space to another in a fixed time interval acts in this manner (Sect. 4.8.5).

4.5 Multi-layer Control Systems

The preceding *sections* presented *multi-lever* control systems whose multiple levers are regulated by the manager to eliminate the distance $E(Y)$ between Y_t and the objective Y^* through an appropriate *strategy*. The systems we have studied so far are *single-layer* ones (Sect. 4.1) in that all the levers have the capability to eliminate the "general" or *first-level* error, $E^1(Y)_t = Y^* - Y_t$. I shall define these levers as *first-order levers*, as they can all be regulated to eliminate $E^1(Y)_t$ (henceforth I shall also indicate the first-level error simply as $E(Y)_t$, since the "general" error is, by definition, a *first-level* error).

There are many cases where the *first-level* error is so large that the action of the *first-order* levers at our disposal is not able to eliminate it, thus leaving behind a residual error I shall define as a *second-level* error: $E^2(Y)_t = 0 - E^1(Y)_t$. This expression clearly indicates that if $E^1(Y)_t$ is permanently different from zero, the "distance" needed to completely eliminate the error is $E^2(Y)_t = 0 - E^1(Y)_t$, a value that represents, in fact, the *second-level* error. In order to eliminate $E^2(Y)_t$ we must activate an *additional* lever, which I define as the *second-order* levers, which are
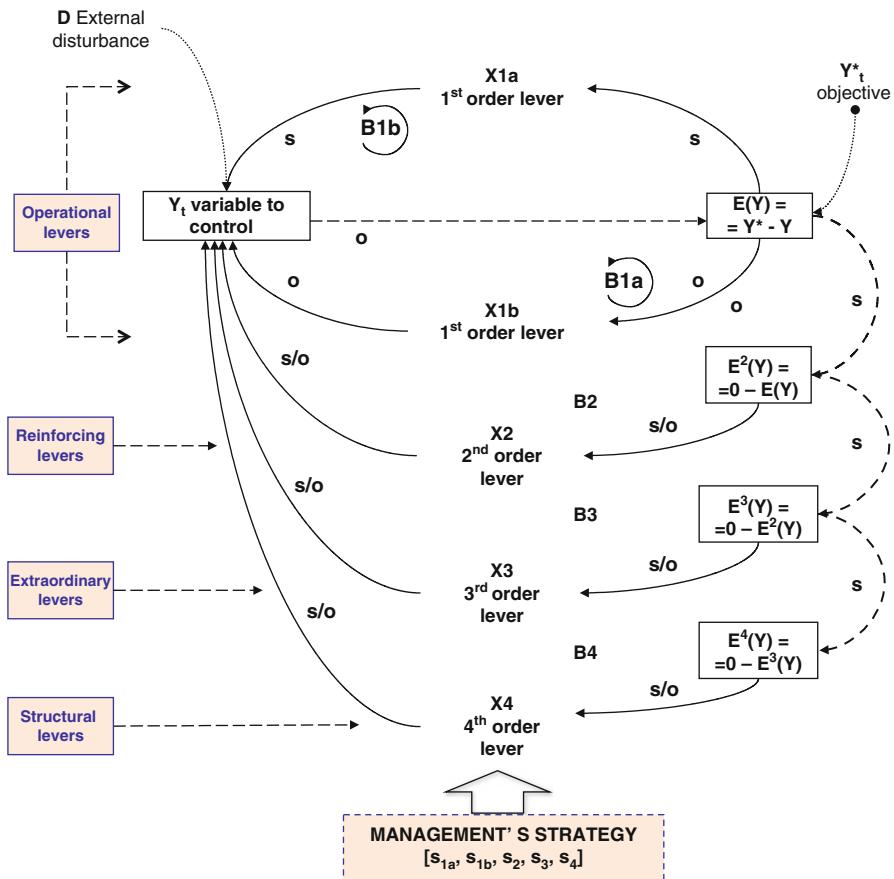
Fig. 4.13   A multi-lever and multilayer control system. General model

added to the *first-order* levers and are regulated only when the latter are insufficient. If there is still a residual *third-level error*, $E^3(Y)_t = 0 - E^2(Y)_t$, after activating the *second-order* levers, then it is necessary to activate *third-order* levers, continuing on in this manner until the residual error is eliminated, if possible.

I shall define a *multilayer* control system as one setup to eliminate the overall error, $E^1(Y)$, by dividing it into different layers, $E^2(Y)_t$, $E^3(Y)_t$, etc., each controlled by levers at successive orders; in this way the control system is divided into various *layers*, each of which use successive order levers to eliminate a portion of the error, as shown in the general model in Fig. 4.13.

Figure 4.13 illustrates a multilayer control model that, though abstract, can be characterized as follows:

(a) The model includes a first-level control obtained through the two action variables, X1a and X1b, which I have assumed to have the opposite dynamic

effect on the variable to control, Y_t; two balancing loops are thus produced: [**B1a**] and [**B1b**].

(b) The model includes successive second-, third-, and fourth-level controls that are activated when management observes that the control at the preceding level is not able to eliminate the error; each error of a certain level derives from the error of the preceding level, when the latter does not reach the desired value of "zero error"; appropriate order levers are activated at each level control.

(c) The controls at the subsequent levels produce the balancing loops [**B2**], [**B3**], and [**B4**]; I have indicated two possible directions of variation, "**s/o**," on the arrows that start from the successive deviations; these arrows must connect with the adjacent ones with direction "**s/o**."

(d) We can define as an *ordinary* or an *operational* control that which is carried out at the *first level* through the *first-order* levers, X1a and X1b, which are *ordinary* levers.

(e) We can define as a *reinforcing* control that which is carried out at the *second level* through the *second-order* lever, X2.

(f) The *third-order* levers, X3, at the *third level*, are usually levers that produce an *extraordinary* control made necessary when the levers from the preceding levels are insufficient.

(g) The levers from the other, successive levels are usually those that modify the structure itself of the apparatuses that produce the control and for this reason are defined as *structural* levers.

(h) There are many possible delays for each control variable (which I have not indicated in Fig. 4.13) whose presence can make the control system complex, in the sense that the behavior of the variable to control can have unpredictable, even chaotic dynamics.

Therefore, in order to define the control strategy the manager must decide to which level the control must be extended, also taking into account the fact that the more the *control level* increases, the more the levers become difficult to activate, thus requiring a longer time to produce their effects on Y_t.

Two immediate examples will clarify the logic of *multilayer* control systems. Several times I have observed that the control of our body temperature is fundamental to our survival. This control can be achieved through a multilayer system similar to that in Fig. 4.14. When we are in a state of normal health there are two physiological levers nature has provided us with: chills and movement help us to raise our body temperature (lever X1a); sweating and resting in the shade allow us to lower it (lever X1b). However, these levers have a *capacity* of action limited to only a few degrees of body temperature, spurring man, beginning in the prehistoric era, to search for strengthening levers in the form of clothing of any kind (lever X2), from primitive furs to modern artificial fibers that nearly completely insulate us from heat and cold. If clothing is not enough, man tries to regulate his body temperature through *extraordinary levers*, so called because they depend on the environment in which he lives and the materials he may find there; these levers include all forms of habitation that insulate him from the outside temperature and
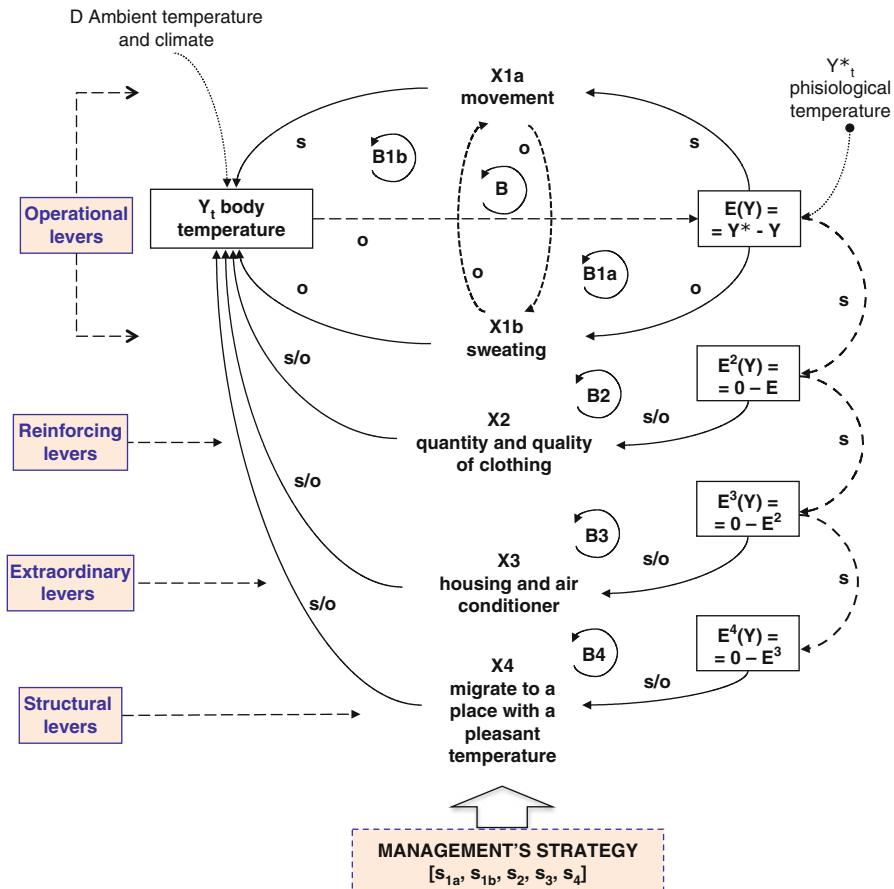
Fig. 4.14 Multilayer control system for body temperature

from the inside one by preserving the artificial heat produced by fireplaces, stoves, and even modern inverter air conditioners for hot or cold air (lever X3). We could stop here at this three-level system; however, recalling the first rule of systems thinking, we could also zoom out further and consider other levers, such as "going to the mountains," "going to live near the poles," and so on.

As a second example, let us consider the familiar system presented in Fig. 4.15: a car (effector) whose driver (manager) wants to control its cruising speed (variable to control) in order to maintain an average speed (objective) or to remain within the speed limit (constraint), even if the steepness of the road as well as other external disturbances, such as traffic and the environment, can alter the speed of the car (Mella 2012, pp. 146–148).

Under normal driving conditions the driver-manager has two main "levers" to control velocity (Y_t): the accelerator (X1a) and the brakes (X1b), which are *linked* and operate in the opposite direction. In fact, the pedals are usually controlled by
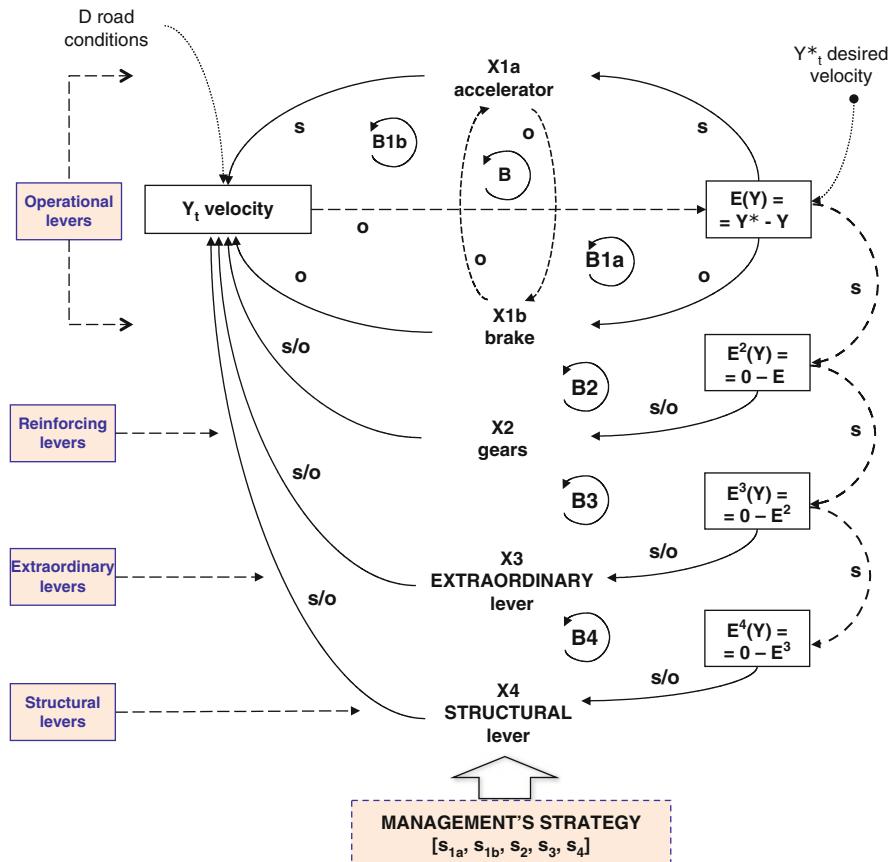
Fig. 4.15 Multilayer control system for car velocity (adapted from Mella 2012, p. 147)

the same foot to make it impossible to use them simultaneously, which in practical terms fixes the relation in direction "**o**." The accelerator and brake "levers" are *first-order operational* levers, since they are regulated by the manager-driver based on the amount of the *first-level* error, $R(V) = Y^* - Y_t$, which is calculated when the speed differs significantly from the objective or the speed limit.

As a *reinforcement* of these two *first-order* levers, the gear shift (X2) can be used along with the action on X1a and X1b as a third operational (reinforcing) lever. Today almost all car manufacturers equip their cars with automatic transmission, which automatically reinforces, without the manager's intervention, the action of the X1a and X1b levers. However, many car models still have manual transmission, which must be activated by the manager-driver, as every driver has his or her own strategy for changing gears. Some use it before the other two levers, while others, to avoid the risk of breaking the engine, prefer to use the gear shift only when they feel that the action of the other two levers is insufficient. However, in general most people choose the strategy of first reducing the pressure on the accelerator lever,

X1a, to zero, then acting on the gear lever, X2 (hitting the brakes), and finally the brake lever, X1b (downshifting). If, when going uphill, the driver sees that the speed has been reduced too much, he or she eliminates the pressure on the brake, steps on the accelerator, and shifts into the appropriate gear. In fact, this is the strategy adopted by car manufacturers for cars with automatic transmissions.

The brake and accelerator are *first-order* levers for a *first-level control*, since they are almost without delay and must be acted on first; the action on the gear shift can be considered as having a delayed effect (even if a slight one), and thus this lever is a *second-order* one for a *second-level* control; this means that the system has an additional *objective* of "zero deviation": $E^2(Y) = 0$. If the action on the accelerator, brake, and gear is not sufficient for the control (considering the time required to reach Y^*), then the manager-driver will detect a further "nonzero error," which conflicts with the second objective of "zero error." There is a *third-level error*, $E^3(Y) \neq 0$, which requires the use of a *third-order* control lever, such as driving zigzag or going onto the shoulder and even rubbing the side of the car against a wall; in extreme cases the driver might even decide to stop the car suddenly by crashing into a tree rather than going headlong into a gulley. These levers can be defined as *extraordinary control levers* in order to distinguish them from the *operational* levers used under normal conditions.

If the control system, through the use of the *extraordinary* levers, permits the car to stop without causing a deadly accident, then the management could modify some parts of the car in order to better control it in future (remember that control systems are recursive), such as by adjusting the brakes, increasing the power of the engine, or changing the gears. It is clear that these control measures can be viewed as *structural* control levers, to distinguish them from *operational* and *extraordinary* control levers, which act on the available levers and not on the structure.

The example reveals a *general conclusion*: in the case of *multilayer* control by means of multiple levers the definition of an order of *priority of actions*, that is, of the *strategy*, for the control lever is crucial and depends on the knowledge (through experience or by referring to the driver's manual) of both the *rate of variation* of each lever—accelerator, brake, and gear shift—with regard to the velocity and the *inverse rate of maneuver* of the levers as a function of the amount of error, taking account of the different *delays* in their "response to the maneuver."

The role of strategy is fundamental in *multilayer* control systems. The strategy affects the efficiency and speed of the control; the experienced manager, following a strategy that has proven successful in past control cycles, is advantaged with respect to one who must choose a control strategy for the first time. This statement is always true, as we can easily verify through observation. The great strategists are the warlords, skilled managers who have gained the necessary knowledge to maneuver their levers, represented by their armed divisions. The great entrepreneurs are skilled managers who know how to maneuver their company's financial, economic, and production levers to achieve success. We ourselves know which levers to use to cook a successful meal, drive in city traffic, find a parking place during rush hour, paginate a book on the computer, or convince someone to see a film and not go to a football match.

I shall define *plan of action*, or *routine* or *consolidated strategy*, as a strategy the manager feels can be effectively applied in all *similar* situations requiring the activation of a control system and which he or she thus *normally* applies in the manual control of all problems requiring analogous control processes. Through *routines* even manual control tends to become transformed into an automatic control, since the manager no longer has to decide on an *ad hoc* strategy each time but rather to carry out a *routine* deemed effective.

Surely none of us has to decide each time which muscles to activate to grab a pencil, to write, to remain in equilibrium, or to walk or which levers to activate to control our car during normal driving conditions. And my wife certainly does not have to establish an ad hoc strategy each day to activate the levers needed to prepare dinner. She is quite capable of cooking even while looking at the TV and monitoring her children's homework, because she follows well-tested *action plans*, well-established *routines*.

The concept of *routine*, that is, a *consolidated strategy*, is mentioned often in studies on organizations, in which a large number of individuals activate control systems to regulate their behavior and the organizational processes without having to select a new strategy each time, instead following well-established routines, procedures, and rules that are adapted to the changing environmental or internal conditions. Routines are referred to by different names. Argyris and Schön, for example, refer to them as "theories of action":

> When someone is asked how he would behave under certain circumstances, the answer he usually gives is his espoused theory of action for that situation. This is the theory of action to which he gives allegiance, and which, upon request, he communicates to others. However the theory that actually governs his actions is his theory-in-use, which may or may not be compatible with his espoused theory; furthermore, the individual may or may not be aware of incompatibility of the two theories (Argyris and Schön 1974, pp. 6–7).

Technical Note: The systems in Figs. 4.14 and 4.15 were created using the same number of levers only in order to better clarify the concept of *multilayer* control system, since it is obvious that systems can exist with a different number of levels as well as levers of the same order.

4.6   Multi-objective Control Systems and Control Policies

All the systems presented in the preceding *sections* have in common the fact they concern a single objective: water temperature or level, velocity, and so on. These limitations are not at all necessary; we can consider even more general models where the *manager* must control several objectives simultaneously by employing an appropriate number of levers with various characteristics.

I shall define as a *multi-objective* (or *pluri-objective*) control system, *in the true sense*, one characterized by a set of objective variables: $Y1, Y2, \ldots, YM$, to be controlled simultaneously in order to achieve an equal number of objectives, under the *condition* that the M variables representing the objective are *interdependent*.

The contemporaneous presence of many objectives does not always signify that the control system is *multi-objective* in the *true sense*; in fact, we can find multi-objective systems which are totally *apparent* or even *improper*. In order to better clarify the behavioral logic of multi-objective systems in the *true sense* it is appropriate to begin with an examination of *apparent* and *improper* multi-objective systems.

Assuming that the M variables, Y1, Y2, ..., YM, though controlled by the same system, are *independent*, each regulated by independent levers that do not interact with the control of other objectives, then the multi-objective system must be considered only *apparent*, since it derives from a simple *operational union* of M single-objective systems. A typical *apparent* pluri-objective system appears during a parade: each soldier that marches must control, at the same time, a multitude of variables: the vertical alignment (with respect to the parading soldiers in front), the horizontal alignment (with respect to the soldiers on each side), the cadenced step, the speed, and, if it is the first row, the marching direction. Even if all these variables must be *jointly* controlled by each soldier in order to have a perfect parade alignment, each of these does not interfere with the others (unless the soldier is distracted) and can be controlled through the appropriate levers by employing different sensors (sight and sound) to eliminate errors with respect to the objectives of a cadenced and aligned march. Each marching soldier can thus be viewed as an autonomous, *apparent*, multi-objective control system. Nevertheless, this system interferes with other control systems, represented by parading companions, and these interferences can involve various control variables. The interferences are evident: for example, when a soldier is "out of step" and causes disorder his fellow soldiers, rather than continuing to march in the same step, try to adjust, causing a series of cyclical oscillations that are clear to everyone that has witnessed a parade. The result is that the soldiers become unaligned and change their marching speed, even only for brief moments, until the control systems (the soldiers) regain control of this variable.

Another *apparent*, multi-objective control system is one we have all experienced when driving in city traffic. Similar to parading soldiers, we have to remain aligned vertically and horizontally by following the flow of traffic surrounding us, which can change speed at any moment. However, in addition we must check the traffic lights, speed limit, route to take to reach our destination, gas level, car temperature, and directional changes, along with many other variables. We realize that managing our car is a difficult task indeed; fortunately, we can control one variable at a time by acting alternately on one and then the other, trusting the speed of our sensors and effectors, but especially the promptness of regulation which experience gives us. Driving slow in city traffic is the sign of a new driver who lacks experience.

A very general case of *apparent*, multi-objective systems is *alternating objectives* control systems, which occur when the manager can alternately achieve different objectives, one at a time, without the momentary objective interfering with the others. As an example, let us assume once again that we are driving our car along a curvy road. The final objective is to keep the car within the driving lane; but to do so we have to, without probably even realizing it, translate this general

objective into two *specific alternating objectives*: maintain a constant distance from the lane dividing line to the left or a constant distance from the right border line. Normally we move from one to the other of the two objectives in continuation depending on the direction of the road. When there is a curve to the left, many prefer to follow the "distance from the lane dividing line" objective, ignoring the distance from the right side; when there is curve to the right the objective focuses on the right border line, not considering the dividing line. Those who try to achieve the *final objective* of staying in the center of the lane by ignoring the specific objectives regarding the distances from the left and right will drive in a swerving manner, as we have all had occasion to observe.

The *alternating objectives* control system is even more evident when we are driving along a narrow road. We need to only recall how often our gaze (detector) shifts to the left or the right of the garage door when we are parking our car. If you need more proof, then consider how our eyes shift continuously when we align letters and lines while writing with a pen, and on the way we pay attention to the two sides of the street when, as pedestrians, we have to cross.

The second type of multi-objective system, which I have defined as *improper*, occurs when the M variables-objectives are *interfering* but we can nevertheless control the achievement of each one; the control system becomes in this case an *improper multi-objective system*, since it is formed by the union of M *interfering* single-objective systems, each of which, however, is able to eliminate the error with regard to its individual objective, unless the interactions among the systems are too strong.

This is the case with a familiar control system: the user-manager who is showering and wants, *at the same time*, to control the water *temperature* and *flow* variables through the use of two levers, the *temperature mixer* (hot–cold) and *faucet*, to regulate the water *flow* (not too strong–very strong). If the mixer can regulate the temperature without affecting the flow, or if the faucet can regulate the flow without interfering with the temperature, then the multi-objective control is only *apparent*, since there is a union of two single-objective systems. The manager must only decide which control variable to use first; in other words, he must establish the priorities for achieving the two objectives (policy). However, if as often occurs the variation in flow also causes a variation in temperature, which will then be corrected using the mixer, then the system can be considered as the *union* of two single-objective *interfering* systems, since the control of flow interferes with the water temperature, even if the reverse is not true (Sect. 2.14).

Another multi-objective system with *interfering objectives* can be observed in an airplane that must fly to a certain destination. The control system represented by the airplane is much more complicated than that of the shower; nevertheless, it has quite similar operating features. The *manager* of this system—the pilot together with the ground staff—must simultaneously control at least three main *operational variables*—direction, speed, and altitude—by using a certain number of levers, as shown in the model in Fig. 4.16. The pilot can *achieve* the final objective, *represented by the flight destination, only by maintaining control of these variables for the entire length of the flight*. I shall leave it to the reader's intuition to interpret the control system in Fig. 4.16, though it is quite clear that what makes the control complicated is the fact that the three variables indicated are interconnected, since
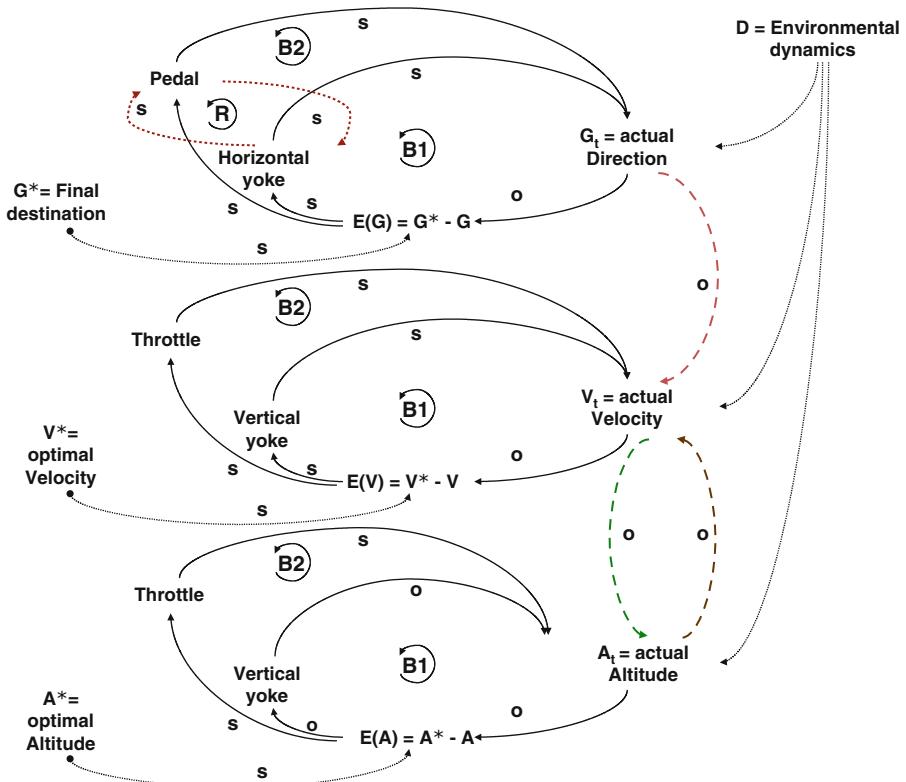
Fig. 4.16 Model of the multi-objective control of a plane

many control levers are common to at least two of them, so that it is not possible to regulate one variable without also regulating the others (see Mella 2012, Sect. 3.7 for a more detailed description).

By observing the rules for the control of speed and altitude indicated in Fig. 4.16, we see that there are no direct constraints among these variables, even though there are indirect ones; this is due to the fact that the *governed* variables are controlled by common commands (levers): an increase in speed cannot be obtained by simply giving the plane the throttle (increase in the number of revolutions of the propeller), since this lever also controls altitude; giving the throttle increases not only velocity but altitude as well. To gain speed the action on the throttle must be accompanied by the maintenance of altitude. Vice versa, by using the vertical control stick to reduce altitude the pilot also increases speed; thus, to avoid this increase he or she must let up on the throttle. In conclusion, the objectives of velocity and altitude are linked by relations in direction "**o**." This is a novelty with respect to the preceding models, since now *the links are placed on the objectives* and not on the control levers of the variables.

Examining these dynamics, we see that the objectives indicated in Fig. 4.16—taking into account the variations adopted by the pilot (manager) during the

trip—are always achieved, even if the reciprocal interactions among these variables necessarily lead to adjustment oscillations. Precisely in order to avoid these oscillations, which would be unpleasant for the passengers, the *designers* of modern aircraft add other control levers to the fundamental ones, which leads to a "precise control" of the *governed* variables: additional ailerons and rudders that automatically intervene to avoid oscillations, control of the engine's revolutions without having to use the throttle, and so on.

In conclusion, even if the objectives interfere with each other the pilot-manager is still able to control the system by bringing all the variables to the value required by the objective*s*. Thus a plane represents an *improper multi-objective* control system in which the pilot-manager always achieves his or her objectives, even if the *interferences* among the variables to control necessarily cause adjustment oscillations.

What makes a system *effectively multi-objective* in a *true sense* is the *condition of interdependence* among the *M* variables to control, so that a control lever regulating a variable Y_i *inevitably* regulates at least one other variable Y_j, whether or not the direction is [**s**] or [**o**]. If both objectives vary in direction [**s**] with respect to the same control *lever*, then the objectives are *cumulative* or *complementary*; if they vary in direction [**o**] they are *concurrent* or *antagonistic*.

It is not difficult to find examples of such systems, even if it is not often easy to describe them effectively. As always I shall begin by proposing a *multi-objective, multi-lever* control system we have all experienced when, as passengers-managers, we have had to regulate the position of the car seat when the position can be controlled through separate *levers* activated by buttons that function by means of electric effectors.

For simplicity's sake I assume that we wish to achieve four objectives: (1) travel safety, (2) passenger-manager comfort, (3) guest passenger comfort, and (4) outside visibility of manager. To achieve these objectives we use three control levers that act on the car seat position by modifying (a) *height*, (b) *distance* from windshield, and (c) *inclination*, according to the following rules:

(1) Safety:

- Low seat
- Seat set all the way back
- Slight inclination

(2) Passenger-manager comfort:

- Seat in middle position
- Seat all the way back
- Strong inclination

(3) Comfort of rear-seat passenger for reasons of courtesy:

- Low seat
- Seat all the way forward
- Strong inclination

(4) Outside visibility:

- High seat
- Seat all the way forward
- Medium inclination

A quick look at this situation immediately makes clear that it is not possible to achieve all the objectives contemporaneously. If we use the lever to move the seat forward and backward we are affecting at the same time safety, manager comfort, passenger comfort, and visibility. Whatever lever we use, all the objectives come into conflict with one another. The system is *truly* multi-objective and the objectives *concurrent*.

Here is a second example. We are preparing dinner for friends and want to control the variety and tastiness of the dishes but at the same time the amount of work needed, time, cost, healthiness of the meal, and so on. It is impossible to totally control all *governed* variables contemporaneously and to reach our objectives at the same time. The control system we are the *manager* of is *multi-objective in the true sense*, since the variables are interdependent and we must prioritize the objectives to optimize. The *governance* of the system must choose, for example, whether to favor quality, by preparing caviar and champagne, or to minimize cost, by serving shrimp and mineral water; whether to cook a pot roast for 12 h or a pasta with sauce in 10 min; whether to use truffles or butter; and whether to worry about cholesterol or flavor the dishes with egg and cheese.

All multi-objective systems, from *apparent* ones with independent levers (shower, car) to *improper* ones with only interfering ones (airplanes), even *multi-objective systems in the true sense*, with interdependent objectives (car seat, dinner), require the manager to ask the *governor* (if the latter is not also the manager) to decide the *order in which the objectives must be achieved*.

When driving over a narrow road with curves the governor-manager must choose if and when to observe the distance from the left and right sides of the road. Under the shower the governor-manager must decide whether to regulate first temperature and then water flow or vice versa. In short, the *governor-manager* would do well to determine an order of importance of the objectives, that is, establish a "control policy" in order to decide which objectives to give priority to. If he does not choose the policy before initiating the control, the policy will be decided "on the field" in a *contingent* manner.

Since controlling a plane toward its destination implies the contemporaneous presence of several partial objectives (direction, velocity, altitude) which must be simultaneously controlled by acting on the specific control variables, it is clear that the *manager*-pilot must decide on a *control policy* that defines the *order* and *priority* of the variables to control and thus the objectives to achieve. In case of a deviation from the route due to lateral wind, he or she could first of all correct the direction and subsequently regulate the altitude or he or she could maintain the actual route and try to reach a better altitude. In the case of headwinds he or she could use the velocity control while maintaining the actual route or he or she could try to maintain the plane's velocity while deviating from his or her route in order to

avoid the headwinds. In the case of tailwinds he or she could maintain the throttle constant to arrive ahead of time or he or she could let up on the throttle to take advantage of the "push" from the wind and arrive on time. Another alternative would be to change his or her route in order to arrive on time without reducing the throttle. The need to decide on a *priority of objectives* in a multi-objective system with *interdependent* and *concurrent* (seat regulation) objectives is much more apparent.

The outline for the general model of control is thus the following: a multi-objective system can be controlled only if the *manager-governor* defines the *set of objectives* to pursue (and the *policy* that determines their relative priority) and if, at the same time, there are sufficient control "levers," which are compatible and coordinated to use based on a defined *strategy*. Several methods for assigning a relative weight of importance, or urgency, to the objectives are presented in Sects. 4.8.8 and 4.8.9.

4.7   Optimal Strategies and Policies: Two General Preference Models

"Control," viewed as a concrete *managerial* action to guide the system toward its objectives, is always an "online" activity, that is, one *concomitant* to the system's dynamics. As such it must be regulated by a *policy* and *strategy* of control the *governor* and *manager* must specify beforehand or while the system is in action. If the *policy* and *strategy* are defined before the system's trajectory begins, the control is defined as *formalized*. If these are decided while the trajectory is unfolding, the control is *contingent*.

The formulation of a control *strategy* implies that the *manager*-user be able to activate the control levers based on an *order of preference* by searching whenever possible for the most efficient and effective order, that is, the *optimal strategy*. The choice of order of preference is not always easy. This problem would require a good deal of treatment, and here I can only present some general, brief indications. I shall start by considering that several *levers* can be activated at no cost (with whatever meaning we wish to attribute to that term, even in the sense of damage, disadvantage, sacrifice, etc.), while others can be activated only with an *explicit* cost (direct sacrifices from activating the lever) or an *implicit* one (lost advantage from activating the other levers).

There is no cost in adjusting the stereo volume; when showering there is no cost in increasing the water temperature by reducing the cold water flow, while there is a greater cost if we increase the flow of hot water (fuel, electricity, etc.); there is no explicit cost to reducing the speed of a car by letting up on the gas pedal (in fact, this reduces fuel consumption), while there is a cost to braking (brake disks are consumed), especially if we take into account the cost due to the damage caused by not braking and having an accident.

In choosing the control levers *cost* plays an important role, though it is not the only parameter to consider. Clearly we must also take account of the *amount of error* and the *importance of the objective*. If there is a speed limit and I note the presence of the highway patrol, I will not undertake an economic calculation of the costs and benefits of braking but try instead to reduce "at all costs" my speed using every available lever, so as to avoid a ticket. Even less would I calculate the costs if my car was going down a steep incline with narrow curves; if my brakes were not sufficient and the speed increased to as to endanger my physical well-being, I would instinctively activate the extraordinary levers (steer into a mound of dirt, scrape the side of the car against a wall, etc.) independently of the cost of activation, even if I know a few people who would rather risk their lives than ruin the body work on their car.

Therefore, the idea of an *optimal strategy* is a relative one and depends on the type of objective and advantages in play and on the amount of error and the costs associated with each lever. When there is time to act rationally by programming the activation of the control levers without being pressured by the urgency of the situation or the gravity of the control, then we can determine the *optimum strategy* through a cost–benefit analysis (CBA), evaluating all the cost factors associated with the levers (direct costs and indirect damage) as well as all the benefits in terms of the advantages from reducing the error and achieving the objective.

In many cases applying CBA techniques encounters considerable difficulties since, in evaluating the different control levers, the manager takes into account monetarily non-quantifiable factors of psychological and emotional origin or factors linked to culture and preferences, which at times are even prejudicial. As proof of this consider the simple control system of the blood pressure of someone who has suffered from "high blood pressure" for some time. To reduce the pressure to its physiological limits, different levers can be activated, among which skipping meals and losing weight, avoiding the use of alcohol, giving up smoking, living a healthier lifestyle, exercising, and taking blood pressure pills. How should a rational person behave in determining an *optimal strategy*? Each of us can provide our own answer, but it is very likely that most people would, irrationally, prefer the blood pressure pill lever to giving up smoking, food, and alcohol (Sect. 5.6).

Similar considerations can be made to define the *optimal policy* of a multi-objective system. The choice of the order to achieve the different objectives depends in part on the costs/benefits associated with the available levers, but even more on the *importance* and *urgency* of the objectives the system must achieve. Usually, the first objective to achieve should be the one deemed most urgent or important, since a high benefit (if achieved) or a high fallout (if not achieved) is associated with it. If the control is not uniquely tied to the urgency of the objectives, then we can add techniques for ordering the objectives on a "pseudo-cardinal scale" to the definition of *optimal policy*, so that the policy choice can entail metrics that assign objective *priority indicators* that are coherent with the idea of the distance between urgency, importance, etc. Several methods have been devised for assigning pseudo-cardinal scales in order to prioritize the objectives of the observed control system.

In *automatic systems* these priority scales are determined by the system's designer and incorporated in the strategy that activates the levers in the most appropriate way to achieve the most important objectives first, followed by the others. For example, a digital reflex camera in the "auto" setting usually prioritizes the objective of optimizing the quantity of light reaching the sensor in order to allow photos to be taken with little light, even if this is at the expense of the objectives of depth of field and still image. In order to allow the manager to better adjust the levers, other operational modes are provided that favor different objectives, which the manager-photographer can choose using simple commands according to the objective deemed most important at a certain moment. A well-known reflex camera provides the following operational modes for the manager to choose based on his or her objectives: "Automatic mode with detection of 32 scenes, steady photos in any situation, gorgeous photos in low light conditions, high-speed, high-resolution photos." Clearly each mode favors a particular objective while ensuring that the other objectives are gradated along a scale of decreasing importance.

In a social context the scale of priorities of objectives is established by norms, customs, or beliefs of various types. Should a control system for saving passengers on a ship privilege saving goods or people? Or saving an adult male, adult woman, a sick person, or a baby? The phrase "priority for women and children" helps to establish the scale of objectives. In setting up control systems for human behavior, the laws of nations and religions often define several scales of objectives that guide the control processes: For example, should we first control tax evasion or privacy? Murder or theft? Theft or corruption? Lying or a code of silence? Prison overcrowding or an increase in crime? Educational levels or resources for schools, etc.?

Scales to order objectives based on urgency, importance, utility, etc. are rarely *quantitative*. Though it is theoretically possible to imagine an objective being doubly urgent or useful compared to another, it is not normally possible to determine or define a *measuring procedure*, since any measurement requires that an objective that can serve as a base unit of a quantitative scale be identified or defined. In order to assign an *index of priority* (urgency, importance, utility, etc.) to each objective we need a procedure that, borrowing a term from topology, I shall define as *metrics procedure*—which differs from measurement—capable of creating *pseudo-cardinal scales*. I shall define *pseudo-cardinal* as a scale whose numbers are not derived from measurement but from procedures which assign *coherent indices* to the various states of a qualitative variable, so that we can compare the dimensions of these states unambiguously by employing the assigned indices.

It is useful to recall two metrics procedures which will be described in Sects. 4.8.8 and 4.8.9:

- The *direct comparison* procedure
- The *linear combination* procedure in the *standard gamble* method variant

However, even the definition of the optimal policy is affected (perhaps to a large extent) by the *personal equation* of the governor-manager, whose scales of

preference reflect cultural factors ("women and children first"), ethical factors ("before food we require safety"), political factors ("first aid to workers then to firms"), and so on. I would also note that the definition of the optimal policy is also affected by several general psychological models which, often unconsciously, influence the choices and limit the rationality of the decision-maker.

It is natural that the rational person has a greater preference for short-term advantages than long-term ones and, conversely, long-term disadvantages are preferred to short-term ones, even if the amount of short-term advantages or disadvantages is less than the respective long-term ones. As we know, these principles are the basis for the financial calculation that provides the techniques for comparing *alternative* financial advantages and disadvantages at different "times."

Nevertheless, Sect. 1.6.6 highlights three systems archetypes (Fig. 1.20) that describe three general structures of human behavior which reveal the widespread tendency for man to behave in an apparent rational manner in an individual and local context, if viewed in the short term, but in an irrational manner if we consider the long-term effects this behavior produces, in other words, at the collective and global level. If we do not apply systems thinking to choose and order the objectives, we run the risk of being "swallowed up" by the short-term archetypes (individual and local preference) and of not properly evaluating the disadvantages this produces in policy formation by the system's governance, with the result that, even though the system permits the objectives to be achieved, this outcome will generate even more serious future problems, which will require even more powerful control systems.

4.8   Complementary Material

4.8.1   Flying in a Hot Air Balloon

We have all imagined how fascinating a flight in a hot air balloon could be, even if most of us have not had this experience. Rising among the clouds and being taken away by the wind must be an extremely pleasant sensation.

Let us try to understand how to take off and land a hot air balloon, that is, how to control the *altitude* of this aircraft. We all know that the upward thrust is supplied by Archimedes' law: *a body receives a push from the bottom upwards equal to the weight of the volume of fluid displaced*. Since hot air is lighter than cold air, in order to overcome the weight of the nylon envelope, the gondola, and the harness, it is necessary for the ball's volume to be sufficiently large to displace an external volume of air capable of supplying the thrust. Figure 4.17 makes it easier to understand the logic of the system to control altitude.
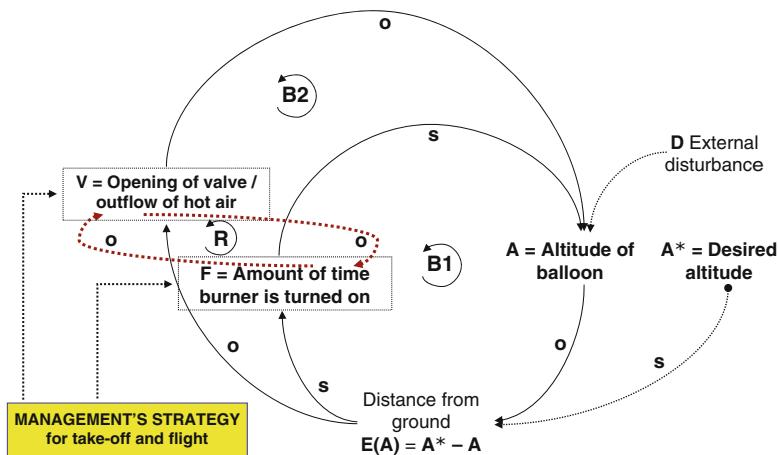
Fig. 4.17 The control system for a hot air balloon

The hot air is not injected into the envelope. A burner—which represents the *control lever* for climbing—in the lower mouth of the envelope heats the air in the balloon; thus, in order to take off the burner needs to only operate for the time needed to heat the internal air mass, based on a heating parameter that depends on the device used. How do we bring it down? It would be enough just to wait for the internal air to cool down, but this usually takes too long to be efficient. There is a *second lever* for landing (the one that controls the descent) represented by a simple valve at the top of the balloon. When opened the hot internal air is released, thereby leaving room for the cooler air that is injected through the lower opening of the balloon. Figure 4.18 shows the general Powersim simulation model for a hot air balloon.

The simulation assumes that initially the balloon is at 100 m of altitude, at which point the pilot ignites the burner to make the aircraft rise to 300 m; after 40 s of flight the pilot orders the air valve to be opened to descend back down to 100 m. After 25 s the balloon lands. The bold line traces its trajectory. The line at the bottom represents the ground level. We can observe that, due to the time needed to heat the air, the ascent does not stop precisely at 300 m but continues on for some meters due to inertia, after which it settles at the desired altitude. For the descent, the adjustment of the valve permits a more precise control, and if the pilot is experienced the airship gently touches down.

4.8.2   Submerging in a Submarine

Even if submerging in a submarine is not as thrilling as a flight in a hot air balloon, it may nevertheless be an equally interesting experience. We are not interested here in the visual sensations but in the control system of this vessel for underwater
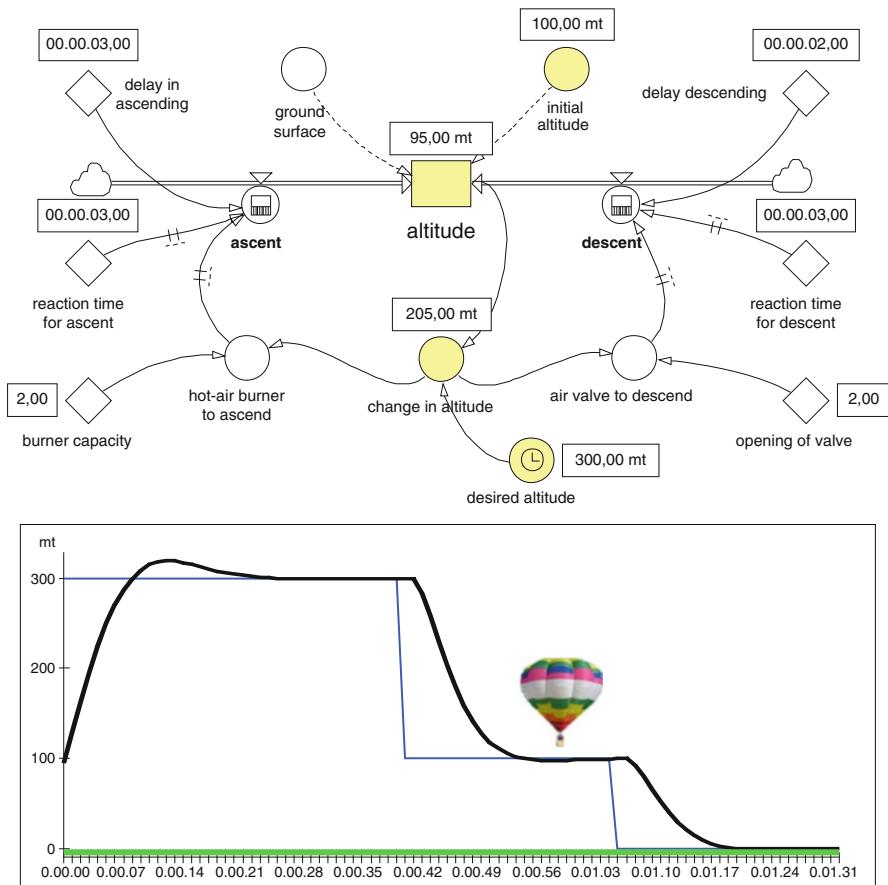
Fig. 4.18 Two-lever control of a hot air balloon, simulated using Powersim

navigation. It should be absolutely clear that the control system for the depth of the submarine—illustrated in Fig. 4.19—is the same as that which regulates the altitude of a hot air balloon in Fig. 4.17. The submarine tends to float as a result of Archimedes' law of buoyancy. It can descend to the deepest depths of the sea only if its mass is increased through the immission of seawater through an open valve that allows the water to gradually fill a certain number of ballast tanks.

When the mass of injected water, added to the weight of the submarine, exceeds the weight of the water it displaces, then the submarine can submerge. In order to surface, a pump injects pressurized air into the ballast tanks; the pressure of air forces the water to flow from the valve. Gradually the total mass of the submarine becomes less than that of the water displaced by its overall volume, and the submarine inexorably rises until it surfaces. In order for the submarine to remain stably at a certain depth it is necessary to adjust its mass—through the immission
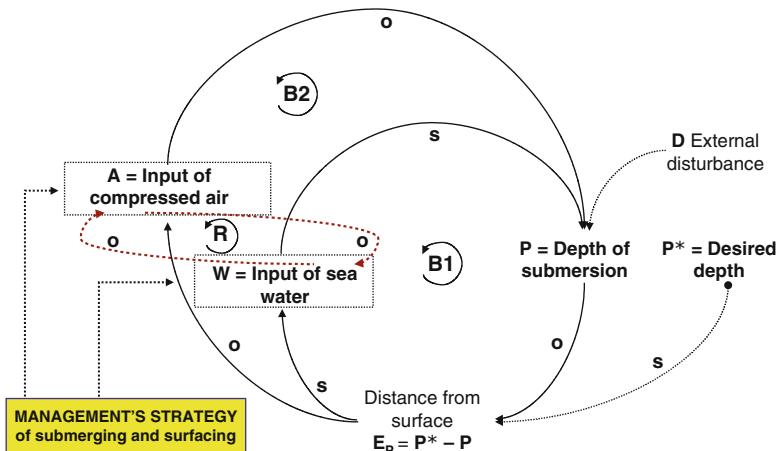
Fig. 4.19 The control system of a submarine

or the expulsion of water from the internal chambers (tanks)—so that it is equal to the mass of displaced water. As in the case of the hot air balloon, the commander-manager of the submarine uses the input control levers (valve for injection of water) and the output ones (compressor for the compressed air used to expel water) in an interdependent way, activating one at a time. Figure 4.20 presents the Powersim simulation of the control system. Initially the submarine is 100 m below sea level and starts to descend to −300 m. After 40 s of descent the commander gives the order to rise back up to 100 m and 25 s after that to resurface. The bold line traces the trajectory.

Note that, due to the time needed to empty tubs and the delays caused by the water's viscosity, the descent does not stop at −300 m but continues due to inertia for some additional meters, after which the submarine rises again and stabilizes at the desired depth. The same dynamics hold for resurfacing. Inertia causes the submarine to exceed the limit ordered by the commander, and when the vessel comes to the surface it jumps out of the water, like a dolphin. The only obvious difference with the hot air balloon is that for the latter the ground represents the lower limit of descent, while for the submarine the upper level of ascent is the sea level. On the other hand, the hot air balloon has an altitude limit that cannot be exceeded; this limit is defined on each occasion by the air density. The submarine has a lower depth limit, which is a function of the vessel's resistance to outside water pressure. The similarities between the two control systems, independently of the differences in the "machines" they are applied to, are immediately clear when we compare Figs. 4.18 and 4.20, which coincide except for the inversion in the direction of motion.
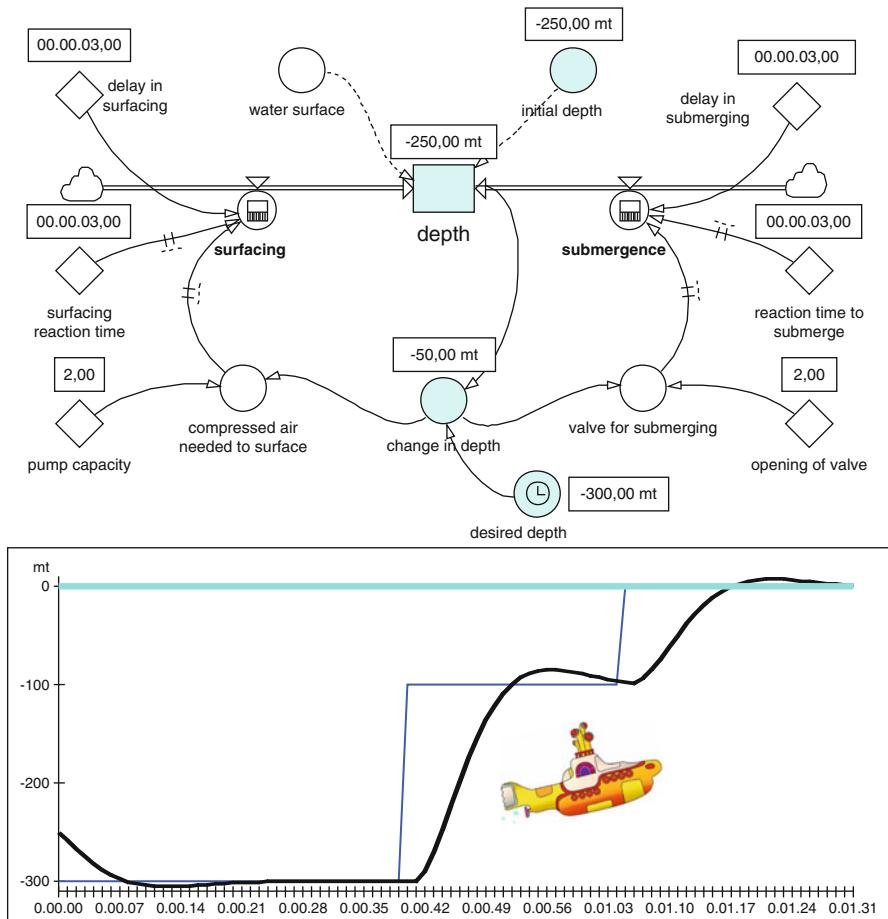
Fig. 4.20 Two-lever control of a submarine, simulated with Powersim

4.8.3 A Multi-lever System: Mix of N Components

Section 4.4 presented an example of the spatial movement of a point in an N-dimensional space, based on the logic presented in Fig. 4.12. A system directly derived from the preceding one enables us to create and maintain a given mix, MIX_t, among N components $[x^1, x^2, x^n, \ldots, x^N]$, with any N (whole, finite). Each component, x_n, must enter the mix in an "ideal" quantity, x^{*n}, which represents the objective; the mix is perfectly formed when all N components have the values $x^n = x^{*n}$. The values of the N quantities making up the mix at instant "t," $MIX_t = [x^1_t, x^2_t, \ldots, x^n_t, \ldots, x^N_t]$, form a point in the N-component space; the control system must shift this point to the point formed by the N quantities-objectives: $M^* = [x^{*1}, x^{*2}, \ldots, x^{*n}, \ldots, x^{*N}]$.

names of legs				X	Y	Z	W	
initial level of legs P_0				3	2	2	1	INITIAL POSITION
height of the roughness of ground P_{ground}				15	12	8	2	15 MAXIMUM HEIGHT
delta extension of legs ΔP_0				-3	1	5	12	
extension of legs P^*				0	3	7	13	OBJECTIVE
				1	1	1	1	r = reaction times
				1	1	1	1	g = action rates = translators

g_x	g_y	g_z	g_w	x	y	z	w	x^*	y^*	z^*	w^*	$dx=x^*-x$	$dy=y^*-y$	$dz=z^*-z$	$dw=w^*-w$	distance
				3	2	2	1	0	3	7	13	-3	1	5	12	13.38
-1	1	1	1	2	3	3	2	0	3	7	13	-2	0	4	11	11.87
-1	0	1	1	1	3	4	3	0	3	7	12	-1	0	3	9	9.54
-1	0	1	1	0	3	5	4	0	3	7	12	0	0	2	8	8.25
0	0	1	1	0	3	6	5	0	3	7	12	0	0	1	7	7.07
0	0	1	1	0	3	7	6	0	3	7	12	0	0	0	6	6.00
0	0	0	1	0	3	7	7	0	3	7	12	0	0	0	5	5.00
0	0	0	1	0	3	7	8	0	3	7	12	0	0	0	4	4.00
0	0	0	1	0	3	7	9	0	3	7	12	0	0	0	3	3.00
0	0	0	1	0	3	7	10	0	3	7	12	0	0	0	2	2.00
0	0	0	1	0	3	7	11	0	3	7	12	0	0	0	1	1.00
0	0	0	1	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00
0	0	0	0	0	3	7	12	0	3	7	12	0	0	0	0	0.00

Fig. 4.21 Horizontal leveling of a platform supported by four legs

The system can act based on constant *action rates* applied over a given time interval or in a manner proportional to the error. These methods of control must be well designed and take into account the nature of the mix, the relative quantities of the components, the reaction times, and the maximum allowable time for completing the control.

4.8.4   Control of a Mobile Platform

The control system to shift a point toward another point, which is described in Sect. 4.4, can be applied in many other ways. One of these is to control the horizontal equilibrium of a platform placed on irregular ground, for example, the platform of a crane, a scaffolding, and a marine platform, where the platform is set on four adjustable legs, moved by hydraulic pistons, which can be set at different heights so as to counter the roughness and unevenness of the ground.

Observing the model in Fig. 4.21, we can describe how the control functions as follows: first of all, the height of the rough ground is measured to form point P_0, and

then the maximum height, P^*, is calculated, that is, the height-objective to which the position of the entire horizontal platform must refer.

Given the initial position of the legs, P_0, the system calculates the necessary extension of each leg with respect to the maximum height, P^*, required to bring the horizontal platform to that maximum height. Referring again to Fig. 4.21, we see from the control panel that, for example, leg X has an initial extension of $x_0 = 3$ units of measure (centimeters, inches, etc.); it must rest at a ground level of 15, which corresponds to the maximum ground height. Thus, the value $x_0 = 3$ must be adjusted and brought to $x^* = 0$ by using a small motor that, for each "t," shifts the leg $g_x = 1$ unit. Leg X is level at $t = 0$. Leg W has an initial position of $w_0 = 1$ and must be level at ground height 15. Taking into consideration that the ground on which W rests already has a height of 2, then leg W must be extended to $w^* = 13$, which, since $g_w = 1$, is reached at $t = 11$. The same procedure can be observed for the other two legs. At time $t = 11$ all the legs are extended so as to level the platform at height 15, thus terminating the control, since the distance from the objective is eliminated (line 11 in the right-hand column). Note that this system also acts to level a mobile platform anchored in another platform, as occurs, for example, with platforms anchored to the ceiling of a cave or a building.

4.8.5   Industrial Robots and Movement Systems

The multi-lever systems in Sect. 4.4 can operate contemporaneously on two *movement systems*, whose dynamics interact according to given rules. Let us first consider the case of two systems, A and B, and apply the rule that A is independent and follows a path toward its spatial objectives, indicated by the coordinates X, Y, and Z, and its temporal ones (coordinate T), while B must follow A and have the objective of reaching, for every "t," A's position, as shown in Fig. 4.22.

A's initial point is $PA_0 = (0, 0, 0)$, which must shift to the objective $PA^* = (40, 40, 40)$. B's initial point is $PB_0 = (40, 40, 40)$, and B must follow A to the distance-objective $PB^* = (30, 30, 30)$. These values indicate the distance each coordinate of B must reach with reference to the respective coordinate of A. The distance of pursuit, calculated from the origin, is 51.96 (placed in the cell positioned between the two systems). This value must be reached by B's trajectory, which follows that of A, and it can be read from the penultimate column of system B (second table). Obviously this value must also be determined by the difference between the distance from the origin of A (last column in the first table) and B (third-to-last column in the second table). As Fig. 4.23 also shows, B's objective of pursuit, PB^*, is achieved at $t = 9$, when the values of the last two columns of system B coincide. Clearly, B is a typical *control system of pursuit*, since B tries to pursue A, whose trajectory represents B's variable objective.

According to the models above, the movement systems with multi-level control are the typical systems studied and produced by *robotics*. In particular, Cartesian robots (which, for example, move a utensil placed at the end of a mobile arm from one point to another in three-dimensional Cartesian space) are referred to as typical

SYSTEM "A"

	X	Y	Z	T	
P = (x, y, z, t) =	0	0	0	0	position
P* = (x*, y*, z*, t*) =	40	40	40	50	objective
action rates-translators	5.0	5.0	5.0	1	

t	gx	gy	gw	gt	x	y	z	t	x*	y*	z*	t*	dx=x*-x	dy=y*-y	dy=y*-y	dt=t*-t	Distance of A from objective	Distance of A from ORIGIN
0					0	0	0	0	40	40	40	50	40	40	40	50	69.28	0.00
1	5	5	5	1	5	5	5	1	40	40	40	50	35	35	35	49	60.62	8.66
2	5	5	5	1	10	10	10	2	40	40	40	50	30	30	30	48	51.96	17.32
3	5	5	5	1	15	15	15	3	40	40	40	50	25	25	25	47	43.30	25.98
4	5	5	5	1	20	20	20	4	40	40	40	50	20	20	20	46	34.64	34.64
5	5	5	5	1	25	25	25	5	40	40	40	50	15	15	15	45	25.98	43.30
6	5	5	5	1	30	30	30	6	40	40	40	50	10	10	10	44	17.32	51.96
7	5	5	5	1	35	35	35	7	40	40	40	50	5	5	5	43	8.66	60.62
8	5	5	5	1	40	40	40	8	40	40	40	50	0	0	0	42	0.00	69.28
9	0	0	0	1	40	40	40	9	40	40	40	50	0	0	0	41	0.00	69.28
10	0	0	0	1	40	40	40	10	40	40	40	50	0	0	0	40	0.00	69.28
11	0	0	0	1	40	40	40	11	40	40	40	50	0	0	0	39	0.00	69.28
12	0	0	0	1	40	40	40	12	40	40	40	50	0	0	0	38	0.00	69.28
13	0	0	0	1	40	40	40	13	40	40	40	50	0	0	0	37	0.00	69.28
14	0	0	0	1	40	40	40	14	40	40	40	50	0	0	0	36	0.00	69.28
15	0	0	0	1	40	40	40	15	40	40	40	50	0	0	0	35	0.00	69.28

SYSTEM "B"

	X	Y	Z	T			
P = (x, y, z, t) =	40	40	40	0	position		
Distance of B from A	30	30	30	50	objective	Distance/pursuit	51.96
action rates-translators	5.0	5.0	5.0	1			

t	gx	gy	gw	gt	x	y	z	t	x*	y*	z*	t*	dx=x*-x	dy=y*-y	dy=y*-y	dt=t*-t	Distance of B from objective	Distance of B from ORIGIN	Distance of B from A	Objective of distance
0					40	40	40	0	0	0	0	50	40	40	40	50	69.28	69.28	69.28	51.96
1	-5	-5	-5	1	35	35	35	1	5	5	5	50	30	30	30	49	51.96	60.62	51.96	51.96
2	0	0	0	1	35	35	35	2	10	10	10	50	25	25	25	48	43.30	60.62	43.30	51.96
3	5	5	5	1	40	40	40	3	15	15	15	50	25	25	25	47	43.30	69.28	43.30	51.96
4	5	5	5	1	45	45	45	4	20	20	20	50	25	25	25	46	43.30	77.94	43.30	51.96
5	5	5	5	1	50	50	50	5	25	25	25	50	25	25	25	45	43.30	86.60	43.30	51.96
6	5	5	5	1	55	55	55	6	30	30	30	50	25	25	25	44	43.30	95.26	43.30	51.96
7	5	5	5	1	60	60	60	7	35	35	35	50	25	25	25	43	43.30	103.92	43.30	51.96
8	5	5	5	1	65	65	65	8	40	40	40	50	25	25	25	42	43.30	112.58	43.30	51.96
9	5	5	5	1	70	70	70	9	40	40	40	50	30	30	30	41	51.96	121.24	51.96	51.96
10	0	0	0	1	70	70	70	10	40	40	40	50	30	30	30	40	51.96	121.24	51.96	51.96
11	0	0	0	1	70	70	70	11	40	40	40	50	30	30	30	39	51.96	121.24	51.96	51.96
12	0	0	0	1	70	70	70	12	40	40	40	50	30	30	30	38	51.96	121.24	51.96	51.96
13	0	0	0	1	70	70	70	13	40	40	40	50	30	30	30	37	51.96	121.24	51.96	51.96
14	0	0	0	1	70	70	70	14	40	40	40	50	30	30	30	36	51.96	121.24	51.96	51.96
15	0	0	0	1	70	70	70	15	40	40	40	50	30	30	30	35	51.96	121.24	51.96	51.96

Fig. 4.22 Pursuit control systems
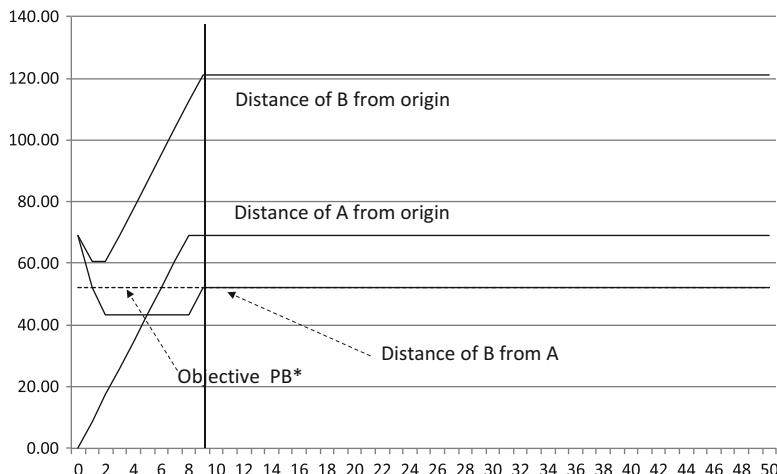
Fig. 4.23 Dynamics of two systems of pursuit in a three-dimensional space

movement systems. Two robots programmed to transfer an object between them (a utensil or a component to work on) can be fully conceived of as *collision* systems. However, if they must place the object at a certain distance from the other they are conceived of as alignment systems.

Cartesian robots can operate in more than three dimensions. For example, let us take an industrial robot that must work on a component in a fixed period of time by carrying out the following program: grab the object; turn it along a vertical axis; choose, one after the other, three tools from its "catalogue"; carry out three successive processes in a fixed amount of time; and, finally, rotate the object and move it forward to permit other objects to be worked on. In this case the robot is in all respects a seven-dimensional (seven-lever) control system: the three spatial dimensions, the speed of the moving objects (shifting time), the degrees of rotation of the transporting head, the manufacturing time, and the catalog of tools.

4.8.6   Focusing

Figure 4.10 presents a simple multi-lever control system that regulates focusing on an object placed on a horizontal plane by rotating the eyes, neck, upper body, and entire body. Figure 4.24 constructs a simple simulation model of this system.

The control panel indicates the action rates for each of the four levers and the maximum range of variation of each lever. These ranges are quantified in geometric degrees with respect to the central focus point, which is equal to "0"; in other words, they are quantified for each semi-plane to the right or the left of the line that is perpendicular to the vertical horizon line. The objective as well is set based on the same logic, with the object to be focused on positioned in the semi-plane to the right or the left of center. I have assumed that the management tries to minimize the time needed to focus on the object and thus adopts the strategy for activating the four levers in parallel, each up to its maximum radius of action.

4.8.7   Demand, Supply, and Price: Dual-Objective Control System

Figure 1.2 in Chap. 1 presented a typical model of the "law of supply and demand," expressed in systems thinking language. According to Walras' classical theory of equilibrium (Walras 1874), the price of a good can be considered a variable controlled by the demand and supply of that good, and not as a control lever. However, Fig. 1.2 shows that, if demand and supply influence price, then price in turn influences demand and supply. If demand rises, price rises; if the price rises then so does supply. Vice versa, if supply falls, price increases, which reduces demand and once again increases supply.

Max Rotation EYES	25	degrees g(X1)	12,5	Position of object	
Max Rotation HEAD	60	degrees g(X2)	5	to focus on	
Max Rotation BUST	90	degrees g(X3)	2	(sin/dx = -/+)	170
Max Rotation BODY	160	degrees g(X3)	1		

t	degrees EYES	degrees HEAD	degrees BUST	degrees BODY	X1 = rotation EYES	X2 = rotation HEAD	X3 = rotation BUST	X3 = rotation BODY	Y = rotation visual focus	Y* =objective (degrees from center)	variation in objective	E(Y)
0	0									170		170
1	1	1	1	1	12,5	5	2	1	20,5	170		149,5
2	1	1	1	1	25	10	4	2	41	170		129
3	0	1	1	1	25	15	6	3	49	170		121
4	0	1	1	1	25	20	8	4	57	170		113
5	0	1	1	1	25	25	10	5	65	170		105
6	0	1	1	1	25	30	12	6	73	170		97
7	0	1	1	1	25	35	14	7	81	170		89
8	0	1	1	1	25	40	16	8	89	170		81
9	0	1	1	1	25	45	18	9	97	170		73
10	0	1	1	1	25	50	20	10	105	170		65
11	0	1	1	1	25	55	22	11	113	170		57
12	0	1	1	1	25	60	24	12	121	170		49
13	0	0	1	1	25	60	26	13	124	170		46
14	0	0	1	1	25	60	28	14	127	170		43
15	0	0	1	1	25	60	30	15	130	170		40
16	0	0	1	1	25	60	32	16	133	170		37
17	0	0	1	1	25	60	34	17	136	170		34
18	0	0	1	1	25	60	36	18	139	170		31
19	0	0	1	1	25	60	38	19	142	170		28
20	0	0	1	1	25	60	40	20	145	170		25
21	0	0	1	1	25	60	42	21	148	170		22
22	0	0	1	1	25	60	44	22	151	170		19
23	0	0	1	1	25	60	46	23	154	170		16
24	0	0	1	1	25	60	48	24	157	170		13
25	0	0	1	1	25	60	50	25	160	170		10
26	0	0	1	1	25	60	52	26	163	170		7
27	0	0	1	1	25	60	54	27	166	170		4
28	0	0	1	1	25	60	56	28	169	170		1
29	0	0	1	1	25	60	58	29	170	170		0

Fig. 4.24 Control of horizontal focus
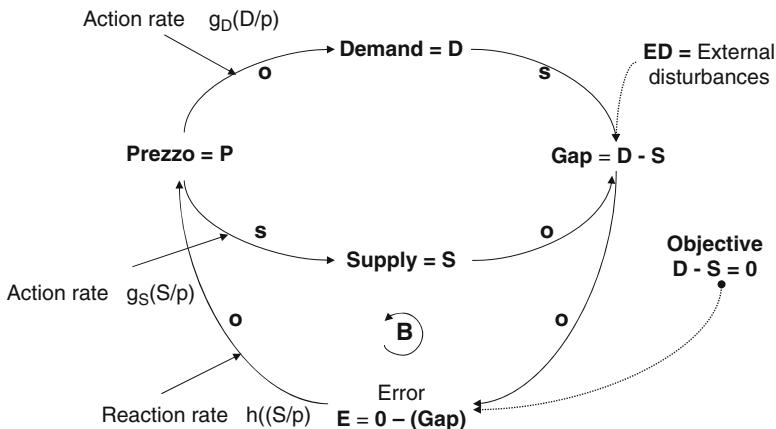
Fig. 4.25 The control system for demand and supply through price

As Mordecai Ezekiel demonstrated in his famous diagram (Ezekiel 1938, pp. 262, 264), which derives from the "cobweb theory," the search for equilibrium between demand and supply and the determination of an equilibrium price can be viewed as the result of a series of re-positionings "on" the demand and supply curves, based on the typical logic of dynamic systems, which, in relation to the slope of those curves, causes prices to converge toward an equilibrium price (in a typical cobweb representation) or to periodically oscillate.

In a simplified attempt at explanation, referring to the model in Fig. 1.2, if $[D - S > 0]$, then the consumers whose demand is left unsatisfied will exercise pressure on the producers, causing them to raise prices in the attempt to increase profits, as they are sure that they will sell their production. If $[D - S < 0]$, then the producers will note an increase in stocks due to lower sales, which will move them to reduce prices in the attempt to eliminate the stocks, thereby providing incentives to consumers to buy their goods.

The model in Fig. 1.2, though interesting from the point of view of a qualitative understanding of the dynamics of the variables, can be improved as soon as we translate it into a control system that transforms the model in Fig. 1.2 into the more efficient one in Fig. 4.25, where the "true" variable to control is the difference between demand and supply $[D - S]$. The model also indicates the action rates "g" and reaction rates "h." The former express the price elasticity of demand and supply. If these rates are constant, then demand and supply have a straight-line trend; if, on the other hand, "g" varies with the size of D and S, then these variables have nonlinear dynamics. The reaction rate could be differentiated for D and S, in which case we shall indicate them as "g_D" and "g_S."

Figure 4.26 presents a very simple teaching simulation model that assumes that:

1. For each unit of price, demand and supply vary by ten units.
2. At $p = 0$ demand is $D = 100$ and supply $O = 0$.
3. For every ten units of gap, price varies by one unit.
4. At $t = 0$ price is $p_0 = 0$.

D_0	Starting Demand	100	g_D = Action rate(price/demand)	10
O_0	Starting Supply	0	g_S = Action rate (price/supply)	10
p_0	Starting price	0	Δ p	1

t	Price	Demand D	Supply S	disturbance D	disturbance S	Error = = (D-S)	Δ p
0	0	100	0			100	1
1	1	90	10			80	1
2	2	80	20			60	1
3	3	70	30			40	1
4	4	60	40			20	1
5	5	50	50			0	0
6	5	60	50	10		10	1
7	6	60	60	10		0	0
8	6	70	60	10		10	1
9	7	70	70	10		0	0
10	7	80	70	10		10	1
11	8	80	80	10		0	0
12	8	90	80	10		10	1
13	9	90	90	10		0	0
14	9	100	90	10		10	1
15	10	100	100	10		0	0
16	10	100	100			0	0
17	10	100	100			0	0
18	10	100	90		-10	10	1
19	11	90	90		-10	0	0
20	11	90	80		-10	10	1
21	12	80	80		-10	0	0
22	12	80	70		-10	10	1
23	13	70	70		-10	0	0
24	13	70	70			0	0
25	13	70	80		10	-10	-1
26	12	80	80		10	0	0
27	12	80	90		10	-10	-1
28	11	90	90		10	0	0
29	11	90	100		10	-10	-1
30	10	100	100		10	0	0
31	10	100	100			0	0
32	10	100	100			0	0

Fig. 4.26 Simulation of a control system of demand and supply (teaching model)

We can observe that, under the simplified conditions assumed in the model, the system attains equilibrium at $t = 5$, when demand and supply are equal, at price $p = 5$. The result is easy to calculate, even intuitively, since the initial difference between D and S is $E = 100$; this causes the price to increase, and since price increases by one unit at each successive instant, after five instants demand falls to $D = 50$ and supply increases to $S = 50$.

To make the example more significant, we assume that for the successive 10 periods after $t = 5$, demand increases by 10 each period ("Disturbance D" column), which is equivalent to an overall increase in demand in the decade of 100 units, so that at $t = 15$ total demand is $D_{15} = 200$ units. Since supply at $t = 0$ is $S_0 = 0$, under the hypothesized conditions the equilibrium price becomes $p_{\text{equilibrium}} = 10$.

At $t = 18$ supply begins to fall over 6 periods ("Disturbance S" column); this is equivalent to a situation where maximum demand is 200, after the increase occurring over 10 periods, while total supply falls to 40. The control system inevitably brings the price to $p = 13$, at which level demand and supply are $D_{23} = S_{23} = 70$.

While we can easily intuit that the price rises to $p = 10$ to equilibrate demand and supply at 100, the increase in price to $p = 13$, with demand and supply in equilibrium at 70, represents a counterintuitive result.

There is no error here; at $t = 18$ the equilibrium $D_{17} = S_{17} = 100$ is altered by the fall in supply by 10 units. The system must readjust to increase price to $p_{19} = 11$ in order to reduce demand and, at the same time, provide the incentive to supply more.

Nevertheless, a new reduction in supply occurs that leads to a further increase in price. This pattern goes on until $t = 23$, when a stable equilibrium is restored at price $p_{23} = 13$. As proof of the correctness in the behavior of the control system of price in Fig. 4.26, we assume that the reduction in supply is eliminated by an annual increase of ten units over six periods ("Disturbance S" column), once again bringing total supply to $S_{29} = 100$. At the end of the decreasing dynamics and adjustment of the system, price returns to $p_{30} = 10$ and equilibrium is again achieved at $D_{30} = S_{30} = 100$.

4.8.8   Ordering of the Objectives to Define Control Policy: The Direct Comparison Procedure

As we observed in Sect. 4.7, in multi-objective control systems the definition of a *policy* assumes the possibility of the *governor* identifying the most "important," "urgent," or "useful" objectives. In order to assign to each objective a *priority index* (urgency, importance, utility, etc.), it is necessary to rely on some *metrics procedure* (which is different from measurement) that can produce *pseudo-cardinal*

scales capable of indicating the "distance" between the values assigned to the objectives. There are two procedures that can be applied:

- The *direct comparison* procedure
- The *linear combination* procedure using the *standard gamble* method

I shall define direct comparison as the procedure which seeks to define the pseudo-cardinal scales by deriving the priority indices through repeated comparisons among the various objectives. This method was originally applied in decision-making theory and in operational research to determine the weights of the objectives or a payoff measure of a given set of predicted results from various courses of action. In this context the method must be attributed to Churchman and Ackoff, in their work entitled *An Approximate Measure of Value* (1954, pp. 172–180).[1] The method is applied not only to the objectives but also to all the objects to which no measurement procedure can be applied but for which the notion of distance or relation between their states makes sense.

In order to understand the logic of this method, which is applied to general objects, an analogy is useful. Let us assume that we have collected five objects, for example, five rocks, and that we want to determine their relative weights in the absence of a measurement procedure. We could first of all order the five rocks according to an estimate of their weight using our sense organs. Let us say that the order goes from O1, the heaviest object, to O5, the lightest. The following question arises: How heavier is O1 than O2, O3, O4, and O5? A scale would have allowed us to give a precise answer, since it *measures* the weight of the different rocks. If O1 weighed 1,000 g, O2 500, O3 400, O4 300, and O5 100, then we could state that O1 is ten times heavier than O5, but also that O1 is twice as heavy as O2 or that O4 is triple the weight of O1, etc.

When it is not possible to measure objects we can turn to the *direct comparison* procedure, which allows us to construct a *relative and approximate* cardinal scale, *ad hoc*, even without being able to measure the objects precisely. The procedure involves three phases:

1. The objects are ordered in a decreasing scale, from the heaviest to the lightest, using some instrument which is easy to construct or simply by using our senses.
2. Each object is assigned a quantitative index, or approximate measure, which, based on our senses, could come close to the weight (which we do not know) of the objects.
3. The first approximation measures are subjected to a series of adjustments so that, at the end of the process, they are consistent with the rational estimate of the differences of weights and relations among the objects; thus, we can state that if O1 is assigned an indicator of 100 and O2 of 50, then O1 is supposed to have double the weight of O2, with respect to the indicators assigned in a coherent

[1] See also Churchman, Ackoff, and Arnoff (1957, pp. 132–139); Ackoff and Sasieni (1968, pp. 51–54), which describes the assumptions for the application of the method. A critical analysis can be found in Fishburn (1964) and in Fishburn (1967, pp. 69–82).

Column 1	Column 2	Column 3
START		
(1) O1 < O2 + O3 + O4 + O5		
(2) O1 < O2 + O3 + O4		
(3) O1 < O2 + O3 + O5		
(4) O1 > O2 + O3 OK		
	(5) O2 < O3 + O4 + O5	
	(6) O2 < O3 + O4	
	(7) O2 > O3 + O5 OK	
		(8) O3 > O4 + O5
		END

Fig. 4.27 Direct comparison procedure. The systematic comparison

manner. The observer must be able to carry out an operational procedure for the ordinal comparison of any subsystems of the objects of observation.

To better understand the above procedure, let us apply it to the five objects mentioned above, assuming obviously that we do not have a graduated scale to weigh them. I suggest that the reader imagine carrying out the metricizing process using a common two pan balance, without, of course, having the weights against which to weigh the objects (for example, 10, 100, 200, 500-g weights), which would have permitted an accurate measurement. We can proceed with the *systematic* and *repeated* comparison of the five objects by placing them in various subsystems in the two dishes.

Let us assume that we have obtained the results shown in the table in Fig. 4.27. The way the comparison works is simple: the heaviest object, O1 (left dish, for example), is compared to the group of remaining objects (right dish). Then O1 is compared to the remaining objects, excluding O5, which is the lightest; then O1 with the remaining objects, excluding O4 as well, and so on until O1 exceeds the weight of the remaining objects in the right-hand dish, which in our example occurs in the fourth comparison (see column 1).

As shown in phase (2), we assume that the five objects have been assigned the *first approximation indices* shown in column [2] of Fig. 4.28. As is immediately clear, apart from the change in scale, the first approximation indices are far from the measured weights of the objects obtained with a graduated scale. A comparison with the table in Fig. 4.27 allows us to verify the adequacy of the indices assigned as a first approximation and to proceed with their gradual adjustment. Let us begin with comparison (8) in column 3 of Fig. 4.27. Since we have to verify that O3 > O4 + O5, then it also must be the case that $n3 > n4 + n5$. However, we instead observe from the indices in column [2] of Fig. 4.28 that $n3 = n4 + n5 = 5$.

Ob-jects	First Approxima-tion Indices	Initial Adjustment	Second Adjustment	Final Adjustment	Definitive Indices	Normalized Indices
[1]	[2]	[3]	[4]	[5]	[6]	[7]
O1	n1 = 8	n1 = 8	n1 = 8	*n1 = 13*	n1 = 13	N1 = 0,43
O2	n2 = 5,5	n2 = 5,5	*n2 = 7*	*n2 = 7*	n2 = 7	N2 = 0,23
O3	n3 = 5	*n3 = 5,5*	n3 = 5,5	n3 = 5,5	n3 = 5,5	N3 = 0,18
O4	n4 = 4	n4 = 4	n4 = 4	n4 = 4	n4 = 4	N4 = 0,13
O5	n5 = 1	n5 = 1	n5 = 1	n5 = 1	n5 = 1	N5 = 0,03
TOT					30,5	1

Fig. 4.28 Direct comparison procedure. Adjustment of the indicators

Thus, it is necessary to adjust the index assigned to O3 to obtain the desired inequality. If we change n3 = 5 to n3 = 5.5, then the indices of O3, O4, and O5 are congruent; they are written in column [3] of Fig. 4.28 [the adjusted datum is in italics].

We then proceed with comparison (7) in Fig.4.27, from which we get O2 > O3 + O5. As a result, we should get n2 > n3 + n5; since we can see from column [3] of Fig. 4.28 that n3 + n5 = 6.5, we see the inconsistency of n2 = 5,5. We can remove this inconsistency, for example, by setting n2 = 7. This new value is written in column (4) of Fig. 4.28, along with that for column (3).

We proceed in the same manner for the analysis of the other comparisons. From comparison (6) we should have n2 < n3 + n4; since, with the indices already present in column [4], the inequality is verified, the adequacy of n2 = 7 is further confirmed.

Comparison (5) should result in n2 < n3 + n4 + n5; since the inequality is confirmed, the adequacy of n2 = 7 is definitively verified (this comparison is superfluous, with the validity of (6) already having been confirmed).

From comparison (4) we should have n1 > n2 + n3; the data from column [4] in Fig. 4.28, where n1 = 8 and n2 + n3 = 12.5, reveals the non-adequacy of n1 = 8. Thus, we could, for example, change n1 = 8 to n1 = 13; this value, along with the preceding ones, appears in column [5] of Fig. 4.28.

Comparison (3) indicates that the index n1 = 13 is consistent; the index n1 > 13.5 would not have been so, however.

The index n1 = 13 in comparison (2) is appropriate; however, n1 > 16.5 would not have been adequate.

The index n1 = 13 in comparison (1) is verified as being consistent.

After the verification of congruence has been completed, the first approximation measures in column [5] of Fig. 4.28 can be considered as definitive, forming a

pseudo-cardinal scale. These measures can then be rewritten in column [6] and normalized, as indicated in column [7], by calculating the ratio of each measure in column [6] to the total for the measures.

We observe that the procedure would have provided equally consistent, though different, results even if the verification process for consistency had started with comparison (1) rather than (8) in Fig. 4.27. However, in this case, as the reader can readily see, based on the numerical data from the example the procedure would have been less efficient, requiring a greater number of "revisions" of the first approximation measures.

The direct comparison procedure is particularly simple and well suited to evaluating the importance or the urgency of the objectives of the multi-objective control system. However, there are a number of limits to the practical application of this method; above all, the governor-manager must allocate the objectives according to an ordinal scale, which is not always possible, especially when there is a circularity of preferences.

Secondly, this technique does not provide measures of importance or urgency but coherent indices assuming a correct observation of the distances. In effect, each measure can be considered as an element of a given interval of uncertainty. It is easy to check that the scale of indicators in Fig. 4.28 would be coherent even if the index $n4 = 4$ rather than $n4 = 5$ is assigned to object O4. Thus the indicators obtained are subjective and are only valid within more or less uncertain boundaries and coherent only when using the judgment scales of the observer.

4.8.9   The Standard Gamble Method

The standard gamble method, or *linear combination procedure* (Sect. 4.8.8), can also be appropriately applied whenever it is necessary to determine the indices of the preferences of objectives or, more generally, of objects with respect to non-quantitative measures. The method is based on the hypothesis that an individual is capable of expressing *coherent preferences* among "linear combinations" of objects evaluated with respect to a qualitative dimension. The method was proposed in this context by John von Neumann and Oskar Morgenstern, in *Theory of Games and Economic Behavior* (1953, p. 26 and following). The authors assembled a system of axioms that guarantee the coherence of the determinations. In particular, the measure for utility must exclude any complementarity or incompatibility among the objects, whether the latter be monetary in nature, objectives, or courses of action.[2]

---

[2] For a general axiomization, see also Luce and Raiffa (1967), p. 23 and following. A complete axiomization in the context of a complex formalization of the method can be found in Champernowne (1969), Vol. 1, p. 9 and following, as well as in Blackwell and Girshich (1954), p. 104 and following.

Assume that there are three objects, O1, O2, and O3, observed in a non-quantitative dimension D (importance, urgency, etc.), and that the observer can order them, for example, in the following way: O1 > O2 > O3.

We also assume that the evaluator-individual can determine a *coefficient of equivalence*, "c," from "0" to "1," so that O2 (intermediate object) is held to be equivalent (~) to the "linear combination" of the other two:

$$O2 \sim cO1 + (1 - c)O3.$$

As the objects have already been ordered as O1 > O2 > O3, it follows that the evaluator-individual:

(a) Holds that if $c = 1$, O2 is then *less* than the linear combination of O1 and O3, with respect to the qualitative dimension D.
(b) Holds that if $c = 0$, O2 is instead *greater* than the linear combination of the other two.

Once "c" has been determined, from *personal choice* or through some *experiment*, the evaluator-individual can assign the following indices to the three objects:

$$n1 = 1, n2 = c, n3 = 0.$$

These indices form a *pseudo-cardinal* scale of order. The measures assigned are coherent and reflect the conditions of observation of the distances among the objects, though this relies on the subjectivity of the observer's choices. This method can also be applied to a number of objects of whatever kind, by combining them in groups of three, if the following conditions, or axioms, apply:

1. Condition of coherence: The observer must be able to order the objects coherently based on an ordinal scale.
2. Condition of combinability: The observer must be able to form linear combinations between any pair of coherently ordered objects.
3. Condition of existence: The observer must be able to determine the coefficient of equivalence for every imaginable linear combination; this is the most difficult condition to find.
4. Condition of comparability: The observer must be able to compare objects and linear combinations and to coherently order them.
5. Composition/decomposition: It is possible to combine several combinations into one or to break up one combination into others based on the elementary rules of algebra.

If the above conditions hold, the linear combination method is equally as useful as the *direct comparison* one in metricizing aspects such as utility or the importance of objectives or objects in any other non-measurable domain.

In fact, the method was originally created to assign measures to the utility of given objects in order to determine the payoff measures in the field of decision-making and game theory. In this context the fundamental problem with the linear combination method (to determine the coefficients of equivalency, "c") is immediately solved by calculating the probabilities the observer (the observer is a decision-making subject) must be capable of quantifying. In this way the linear

combinations become "lotteries," or "standard games," that involve the various objects; in this case the objects were payoffs deriving from the carrying out of courses of action or strategies. For this reason, in the context of decision-making and games theory the procedure is also referred to as the *standard gamble method*.

The method can be applied with several variations to illustrate which I shall turn to a simple numerical example applied in decision-making theory. If a decision-maker, in hypothetically carrying out four alternatives of action, expects to obtain the following results, R1 > R2 > R3 > R4, which follow a decreasing order in terms of preferences, we can also assume that the results are objectives of a control system or represent other non-measurable objects of observation. We need to assume that the decision-maker can create "lotteries" where those results must be obtained with the given probabilities "p," which play the role of the coefficient of equivalence "c" of the general method.

Let us consider a *first variant* of the *standard gamble* in which the decision-maker creates the following pairs of lotteries:

LOTTERY A: [R1 with probability p1 + R4 with probability (1 − p1)] or [R2 certain]
LOTTERY B: [R1 with probability p′1 + R4 with probability (1 − p′1)] or [R3 certain]

Assume that the decision-maker is "forced" to choose between the lotteries (which correspond to the linear combinations among the objects of observation) and is indifferent in his or her evaluation of the two terms of LOTTERY A, if $p = 30/45$, and the two terms of LOTTERY B, if $p′ = 15/45$. These probability valuations give us the following preference indices:

$$n1 = 10, n2 = 6.7, n3 = 3.3, n4 = 0.$$

These indices are coherent with the observation of the distances among the preferences toward the various results, even if result n4 is difficult to manage with a ratio scale, thereby making the method cumbersome in many situations.

Now let us consider a *second variant* of the *standard gamble* which overcomes this limitation; in this case the decision-maker has to choose among three lotteries (there are four results):

LOTTERY A: [R1 with probability p1] or [R2 certain]
LOTTERY B: [R1 with probability p′1] or [R3 certain]
LOTTERY C: [R1 with probability p″1] or [R4 certain]

It is assumed now that the decision-maker, forced to choose among the lotteries, evaluates indifferently the two terms in LOTTERY A, if p1 = 80 %, the two terms in LOTTERY B, if $p′ = 50$ %, and the two terms in LOTTERY C, if p″1 = 15 %. Using R1 as the term of comparison, we can then assign the following preference indices:

$$n1 = 10, n2 = 8, n3 = 5, n4 = 1.5.$$

These indices do not coincide with those assigned in the first variant of the *standard gamble*, since the lotteries whose indifference probabilities have to be

evaluated have changed; however, in this case as well the method guarantees the coherence of the resulting scale.

There is a *third variant* of the *standard gamble* we can consider, in which, in place of the preceding lotteries, the decision-maker is faced with the following ones, whose main term of comparison is R4 instead of R1:

LOTTERY A: [R1 with probability p1] or [R4 certain]

LOTTERY B: [R2 with probability p2] or [R4 certain]

LOTTERY C: [R3 with probability p3] or [R4 certain]

Assuming that the decision-maker, forced to evaluate the lotteries, has chosen the following probabilities, p1 = 19/200, p2 = 1/8, and p3 = 2/7, we can assign the following preference indices:

$$n1 = 200/19 \approx 10.5, n2 = 8.0, n3 = 3.5, n4 = 1.$$

This third variant appears, at least in theory, to be preferable since, like the *direct comparison* method, it takes as its term of comparison, or *scale unit*, the "lesser" object (R4, in our example).

The limitations of the *linear combination* procedure are evident: first and foremost, the assumption of the comparability of the linear combinations and, secondly, the choice of the *coefficient* of indifference. For this reason, the *direct comparison* method is perhaps more "operational," as this method involves only comparisons among the objects and aggregations of objects (sums, non-combinations of the dimensional states).

4.9 Summary

We have learned that:

1. Systems can be multi-lever (Sect. 4.4), multilayer (Sect. 4.5), and multi-objective (Sect. 4.6).
2. The simplest multi-lever systems are the two-levered ones: one lever, X1, to produce increases (decreases) in *Y*, and another, X2, to produce variations of equal or opposite sign in *Y*. The levers can be:

    (a) *Mutually dependent* (Sect. 4.1), or linked, if X1 and X2 can only produce variations of the opposite sign as *Y* (Fig. 4.1).
    (b) *Independent* (Sect. 4.2), or free, if both X1 and X2 can produce variations of the same sign as *Y* or variations of the opposite sign (Fig. 4.3).

3. A particular type of two-lever system is an impulse system (Sect. 4.3), typically used to control warehouses and biological life; such systems can be:

    (a) CI–PO functioning (Figs. 4.7 and 4.8).
    (b) PI–CO functioning (Fig. 4.9).

4. Multilayer control systems (Sect. 4.5) can have levers of various orders that act at different levels of control (Fig. 4.13), specifically:

   (a) *Operational*, or first-order, levers; these are used immediately (they can be free or linked).
   (b) *Reinforcing*, or second-order, levers if they are complementary to the operational levers.
   (c) *Extraordinary*, or third-order, levers; these are used when the first-level levers are insufficient to eliminate the variance.
   (d) *Structural*, or fourth-order, levers; these are used even if the extraordinary levers are not sufficient to achieve the objective.

5. In multi-lever systems the manager must determine the order of activation of the levers, that is, the *control strategy* (Sect. 4.5), by carrying out a CBA to identify the most effective and efficient levers (Sect. 4.7).

6. The multi-objective systems (Sect. 4.6) are those where the manager simultaneously controls different objectives by employing levers of varying levels (Fig. 4.16); these can be of three types:

   (a) *Apparent*, if they can be conceived of as the conjunction of several independent single-objective systems.
   (b) *Improper*, if they derive from various interfering single-objective systems.
   (c) *Proper*, if they must achieve multiple interdependent objectives.

7. In multi-objective systems the manager must define the order of priority of the objectives, that is, the control *policy* (Sect. 4.7), which must be linked to the control *strategy*, which as we have seen concerns the choice of the control levers.

8. The definition of the control policy must occur on the basis of accurate techniques to order the objectives, which involve defining their scale of priorities, according to metricization procedures that determine pseudo-cardinal scales (Sects. 4.8.8 and 4.8.9).

Chapter 5
The Ring. Observation and Design

Many workers in the biological sciences—physiologists, psychologists, sociologists—are interested in cybernetics and would like to apply its methods and techniques to their own speciality. Many have, however, been prevented from taking up the subject by an impression that its use must be preceded by a long study of electronics and advanced pure mathematics; for they have formed the impression that cybernetics and these subjects are inseparable.

The author is convinced, however, that this impression is false. The basic ideas of cybernetics can be treated without reference to electronics, and they are fundamentally simple; so although advanced techniques may be necessary for advanced applications, a great deal can be done, especially in the biological sciences, by the use of quite simple techniques, provided they are used with a clear and deep understanding of the principles involved (Ross Ashby 1957, Preface, p. v).

The previous chapter completed the presentation of the logical structure of control systems by introducing two important generalizations: (1) *multi-lever* control systems and the concept of *control strategy*; (2) *multi-objective* control systems and the concept of *control policy*. The chapter also mentioned the concepts of *cost–benefit analysis* and *scale of priorities* as basic elements in determining control strategies and policies. In this chapter, which generalizes the considerations presented in the preceding ones, I shall outline the guidelines for recognizing, observing, or designing control systems and the problems that arise regarding their logical realization. I shall introduce the fundamental distinction between symptomatic and structural control. A symptomatic control is short-term and tends to eliminate the error by acting only on the error itself through symptomatic levers. A structural control acts on the effective causal structure between X and Y; it is long term and tries to permanently eliminate the error by acting on Y. I shall then examine the concepts of the efficiency and effectiveness of control systems and the various techniques for increasing the effectiveness of control through the use of internal and external interventions to strengthen the system. A review of the various possible risks of failure of the control process due to structural causes will then be presented, with particular reference to failure caused by the use of the wrong levers. Two fundamental archetypes, "Fixes that fail" and "Shifting the Burden," will clarify how these risks manifest themselves in a widespread and recurring manner. The last part of the

chapter considers the *human aspects of control*, which involves several ontological aspects of control: discouragement, insatiability, persistence, and underestimation. The chapter concludes by presenting the logic behind the decision-making process or behind the replacement of the present states for desired future ones; the decision-making process can, in all respects, be viewed as a unique control process.

5.1   How to Recognize or Design the Logical Structure of a Control System

After presenting the theory of *multi-lever* and *multi-objective* control systems, and in order to be prepared to recognize the vastness of the array of control systems in all observational and operational contexts which will be examined in Part II, the reader will have the opportunity to apply himself in the control discipline. In order to fully achieve the objective set out in Sect. 1.1—to systematize the control discipline and favor its application and spread—it is necessary to complete the task by proposing several guidelines for the correct *observation*, analysis or *design* of the *logical* structure of control systems. It is useful to recall in this section some of the many aspects and problems we have encountered in the analysis of control systems in previous chapters.

We must always keep in mind that, even before we realize the *physical* structure of the system, *observing* or *designing* a control system means understanding or designing the *logical* structure of the control process. The first step in configuring the *logical* structure of a control system is to identify the *governed* variable Y (or variables, in the case of multi-objective systems) to control. This is not a free choice, since obviously the control is necessary when the manager of the system identifies or sets an objective or constraint, Y^*. The *objectives* or *constraints* determine the choice of the variable whose dynamics allows the objectives or constraints to be achieved.

In the examples illustrated in the preceding sections, many control systems were prearranged for the achievement of an objective represented by a specific value of the variable Y_t. We define these objectives as *goal* or *punctual* objectives. Many control systems have a more complex functioning, since they must produce a dynamics in Y_t that conforms to a succession of differently timed values. A succession of values of Y_t that the control system must achieve is defined as a *path* or *tendential objective*, and the system is a *tracking* control system. It is one thing to reach a destination—for example, a city street—and another to do so by taking a certain route with constraints, for example, passing through the center or by your aunt's house. This second type of objective is more difficult to identify, and often the observer views it as a simple succession of punctual objectives or as variations of an original *goal* objective.

Objectives can be further broken down into *static* (or fixed) objectives, which have a sole value (goal)—or a fixed sequence of values (tendential objectives)—to achieve, or as *dynamic* objectives, which vary according to a function that depends

on the values of the variable Y and/or X, as analytically illustrated in Sect. 3.4, which can be applied to the case of two systems.

It is obvious that a hungry lion has the objective of reaching the gazelle that is calmly nibbling the grass in the savanna; or, better yet, to reach the point where the gazelle is at that moment. If the lion is downwind and the gazelle is relatively still while nibbling, then its position represents a *static goal* for the approaching lion, who can easily activate a system of collision (Sect. 4.8.5). Since it is well known that lions do not prefer that tactic but instead approach their prey trusting in their speed and strength, the *position* of the gazelle, as we can easily imagine, almost never represents a static objective. The gazelle, perceiving the danger from the predator, starts to flee, so that its position becomes a *dynamic* objective that the pursuing lion's trajectory must reach.

This *dynamic* objective—the gazelle's trajectory Y^*—that the lion's trajectory must reach can be considered a function of the lion's trajectory, Y_t, and thus its exertions, X_t, as well as of the gazelle's position in the chase environment, Y^*_t; thus we can more generally write: $Y^*_{t+1} = F(Y_t, X_t, Y^*_t)$, using one of the forms presented in Sect. 3.4. On the other hand, the position of the flower the bee wants to reach represents a clear example of a *static* objective. We realize that in the struggle for existence all prey are the objectives of predators; for vegetarian predators, the prey (their position) generally represents a *static* objective; for carnivorous predators the position of a live prey is vital to reach and normally represents a *dynamic* objective. We must recognize that these two types of objective are present in all aspects of life, even the sporting one. In fact, many interesting control systems in the world of sports have variable objectives. Though I enjoy watching the high jump, long jump, the running and walking competitions, the field events with their many types of equipment, and cycling events (all sports which have a fixed objective), I find it much more pleasing to watch baseball, rugby or soccer, all sports where each player is obviously a control system trying to achieve a variable objective, with variations that depend in part on external disturbances and in part on actions by the player himself.

Many objectives are *quantitative* and represent desired numerical values for specific quantitative variables, which can be measured using a given procedure which is defined from time to time. Along with these we should not underestimate the broad category of *qualitative* objectives (flavors, colors, forms, etc.) set by the manager-governor based on his needs and experience.

The producers of food and drinks (wines, cheeses, cookies, soft drinks, etc.) or those that prepare them (chefs, restaurateurs, etc.) well know how difficult it is to maintain a long-standing flavor standard, even with changes in the quality of the raw materials, and for this reason they set up sophisticated control systems to guarantee constant quality over time.

Artists also "know" how their work will look in the end before producing it, and this knowledge guides their work toward achieving their vision (achievement); the great Michelangelo clearly expressed this idea when he said that *the statue already exists in the block of marble, and that the sculptor only has to free it* by eliminating the excess marble. A quote normally attributed to Leonardo clarifies this concept.

In every block of marble I see a statue as plain as though it stood before me, shaped and perfect in attitude and action. I have only to hew away the rough walls that imprison the lovely apparition to reveal it to the other eyes as mine see it (Shaikh and Leonard-Amodeo 2005).

Perhaps most of the control systems in operation every day, and which allow us to have a normal existence, are those characterized by qualitative objectives. We ourselves change restaurants or cafés when we feel that the qualitative objectives of flavor, courtesy, and environment are not met. Similarly, my mother-in-law knows when the wash is not perfectly clean when it comes out of the machine, just as my mechanic knows that the noise from my motorcycle's engine is different from what one would expect.

The presence of qualitative objectives may make the formalization of control systems more complicated, but this in no way changes their logic. The most difficult *qualitative* objectives are forms (faces, environments, models, fingerprints, stripes on a blanket, etc.), which all animals equipped with sight, including humans, learn to recognize better from repetitions of the recognition cycle, time after time, which allows them to construct a catalogue of forms. This process helps them to live and survive.

As *recognizing systems* (Sect. 3.2) we recognize the face, voice, posture, walk and handwriting of an incredibly high number of people and animals. We recognize the attitude, gestures, clothing of friends and enemies, just as animals recognize the odors, footprints, plumage, etc., of friends and enemies. In many cases it is not always easy to identify the control objectives and variables, since a variable to control can sometimes be only apparent. This occurs whenever Y_t is part of a causal chain (for example, A–B–C–Y). Since Y varies as a result of the variations of the variables preceding it in the causal chain, the "true" variable to control may not be Y but A, making it then necessary to identify an objective to assign to A, say A^*, which is entirely equivalent to the objective Y^* we desire for Y_t, so that when A reaches A^*, Y reaches Y^*. It may be possible to control Y with ad hoc levers, but this control would only be temporary since the subsequent variations in A would take Y to non-optimal values. I shall define this situation as *apparent control*, since there is an attempt to constrain variable Y without, however, controlling the variable A, from which Y's dynamics structurally derive.

I would like now to consider a new question: how many levers must we consider in order to achieve complete recognition or to construct a meaningful control system? If we go back to the example of the two-lever control system of water temperature in Fig. 4.3, we immediately realize that in the case of a control with two independent levers, the variable to control (the water temperature) derives, in fact, from a *mix* of hot and cold water. What our skin perceives (sensor apparatus) as a symptom is not so much the *volume* of hot or cold water but the sum of the temperatures of the two water flows, obviously weighted based on the size of the flows. Thus, it is possible to maintain this ratio by varying both the temperature of the flows and their size.

In our showers we do not have the problem of varying the temperature of the water that flows from the two taps (or the problem arises only in the event of outside

disturbances) since usually the temperature of the hot and cold water is maintained constant and the perceived temperature—from the *mix* of the two flows—depends only on the size of the two flows, given the constant water temperatures. If it were possible to vary both the size and temperature of the flows, then the control system of the shower would become a *four-lever system*. Since it is much more efficient to regulate water temperature using a mixer, as in the model in Fig. 2.7, showers with two separate taps have almost disappeared.

We can generalize: if in many cases control systems can be *observed* or *designed* with a *few* or with *many* control levers, then the *observers* or *designers* are torn between two tendencies: *on the one hand*, whether or not to increase the number of control *levers* [Xn], in order to allow a "precise control" of the *variable Y*, and on the other whether or not to reduce the number of *levers* [Xn], through appropriate *mixes* of their values using special apparatuses, in order to make the "control quick" and easy for the management of the system.

How many of us would be content to take a shower in a multi-lever control system where you had to control both the water flow and temperature? That is why modern mixers today regulate the shower very quickly and more comfortably. A chemical plant where the regulation of the mixture temperatures must be quite precise is another thing altogether. How many of the readers remember all the *manual* regulations that were necessary to take a good photograph with the "old" reflex cameras? How many of us hurried to replace our cameras with one that automatically regulated light and focus? How many of us who now have digital cameras that are completely "automatic" still remember the problems we had with manual settings for the shutter opening and shutter speed?

How should the observers and designers of control systems *behave*? Should they increase or decrease the control levers? Should they search for "precision" or instead aim for the "speed and simplicity" of regulation? We can only answer in general: observers and designers must above all determine the function the control system must carry out and whether or not the system is part of a larger system; subsequently they must evaluate whether or not to *zoom out*, toward a more general perspective (by reducing the number of control levers with appropriate *mixes* of action variables) or *in*, taking in more detail (by increasing the number of [Xn]).

The above observations pose a more general question: are there some criteria that suggest how many objectives [Ym] to control simultaneously and how to identify the optimal number of control levers [Xn]? Recalling the general rules of systems thinking proposed in Sect. 1.2, in particular the *first* (zoom in and out) and the *fifth* (specify the boundaries), the answer is a number of jointly controlled objectives and a number of control levers jointly considered depend on how the manager of the control observes reality.

We must always keep in mind (Mella 2012, p. 101) that the models that represent control systems must also (or above all) determine the essential aspects of the logical reality they represent, and thus must identify and connect all, and only those, levers [Xn] which the observer holds to be truly significant for his control objectives [Ym]. The more we zoom in "toward the smallest details," the more the number of levers are reduced; the more we zoom out "toward the larger picture," broadening the

observed environment, the more the control systems will involve multiple objectives and imply multiple levers.

The eye system alone (zooming in on a narrower perspective) can automatically control the act of focusing on various objects (Fig. 4.10), even in movement, through the process of "accommodation"; that is, the rapid contraction of the crystalline lens (effector) using the ciliary muscle (regulator) which, by adding diopters for close-in focusing, guarantees a clear image (focus) even for objects very close to the eye. There are various disturbances to the focusing capacity, the most common of which are, as we know, myopia (the difficulty in focusing on distant objects), presbyopia (which inhibits our ability to focus on nearby objects), and hypermetropia (which makes it difficult to focus on both distant and nearby objects). An auxiliary lever that facilitates the ability to focus on objects is an increase in sunlight, natural, or artificial. If we zoom out and also consider the individual during his lifetime, we can add a third control lever: eyeglasses or contact lenses, which adjust our focus artificially. There are also structural levers that correct visual disturbances by intervening on the structure of the eye or by using excimer laser treatment or other surgical interventions deemed appropriate for more serious cases.

We have reached a *general conclusion*: the number of objectives and control levers of any specific control system does not depend so much on the "nature" of the system as on the extension of the *boundary* we place to circumscribe the reality we are observing or designing. In other words, every control system—and the model built to represent it—is not absolute and objective but relative and subjective. We can further *generalize*: once we have set up a control system, we can usually set new objectives and add other levers by zooming out to get "a broader picture," considering that an increase in the number of objectives to achieve also means increasing the number of control levers and the number of layers needed to achieve full control.

5.2   Symptomatic and Structural Control

After having recognized (observation) or defined (designed) the variable to be controlled, Y_t, the objective to attain, Y^*, and the control lever, X_t, it is important to reflect on the significance of the error $E(Y)$ (I am referring to single-lever and single-objective systems, but all of the following considerations hold for any more general control system). In formal terms, for the control system as such this variable has no other significance than that of representing the *distance* (gap) from Y^* for each instant of detection of the values of Y_t. For the system *manager* this variable has, however, a completely different meaning, which is connected to the manager's physical shape and experiences. In many cases $E(Y)$ is identified by the manager–user as a *symptom* that is physically associated with distance.

In the shower, $E(Y)$ indicates if the water temperature is below or above the desired temperature, while for the manager–user it signifies the sensation of cold, hot, or well-being. For the manager–user of the radio system, $E(Y)$ signifies deafening noise or imperceptible whispering. In the system that controls the

human body's state of nutrition or hydration, $E(Y)$ represents the sensation of hunger, thirst, or satiety. Tiredness is the physiological sensation that derives from the difference between the tolerated level of lactic acid in the muscles and the actual level.

We need to get accustomed to the fact that almost all our sensations of *need* or *satisfaction*, of *dissatisfaction* or *satiety*, can be interpreted as the meaning we attribute to the perception of a deviation between a normal physiological or mental state Y^* (objective or limit) and an actual state Y. They are the *symptoms* of disequilibrium between the actual situation and normality (objective or limit). These considerations are important when we face the next step in identifying the control levers that enable Y to reach Y^*, thereby eliminating $E(Y)$.

The search for the control levers must meet the general requirement of each *Causal Loop Diagram*; the X must have a causal relationship with the Y. Before making any further considerations, and going back to what we have just observed about the meaning of the error, we can state that, rather than finding a causal link with Y, the observer or designer often identifies an immediate link between X and E (Y). For example, thirst, as we see in Fig. 4.6, is not a symptom of the lack of liquid but of the alteration of the salt/water ratio with respect to a normal value; it is a value of $E(Y)$ and not of Y. Thus, it should be clear that the causal link is between the water intake and the salt/water ratio. It would be a mistake to see a direct causal relation between the amount of water drunk and the intensity of thirst, which could lead to catastrophic consequences for our organism.

The same is true for pain, which in general does not represent the variable to control but the *symptom* of a difference (error) between a normal and an altered state of some part or organ in our body. Yet many of us think that pain is cured with analgesics. It is not pharmacologically wrong to take a painkiller for a headache, but the cause of the headache is to be found in some physiological alteration (stress, indigestion, cold, etc.); the control levers for a headache must try to eliminate this alteration, not only the symptom; X must act on Y and not on $E(Y)$.

However, in many circumstances when we cannot identify the lever that acts directly on Y—since the causal link between Y and its causes is unknown—we have no choice but to settle for acting on $E(Y)$, as if this were the true variable to control. For that matter, the curative power of some herbs for a headache has been known for centuries, even if it is only today that we know about many of the *levers* to use to eliminate the true physiological causes of headaches, which act on the chemistry of the brain or on the stress produced from the action of other organs (digestion, liver, etc.). Unfortunately not all *levers* that could control headaches are known, and the spread of headache centers is proof of this.

These considerations allow us to understand the distinction between *symptomatic* and *structural* control systems. Taking doses of analgesics to keep a headache under control, just like spreading on cream to fight muscle pain after a workout, eating to satisfy our hunger, or a digestive against overeating are merely attempts to find the *levers* that produce a *symptomatic* control.

A *symptomatic* control is short term; it eliminates the error by acting only on the error itself by using *symptomatic levers* not on its causes. In subsequent repetitions

of the cycle the error is inevitably destined to return, since its causes have not been eliminated. On the other hand, a *structural control*, precisely because it acts on the effective *causal structure* between X and Y, is long term and tries to permanently eliminate the deviation by acting on Y.

Therefore, a *rule* to follow in the discipline of control is the following: *in observing or designing control systems we must identify the levers that can intervene on the structural variables that produce the symptoms, and not only on those acting on the symptoms*; in other words, we must identify a link between the *structural levers X* and *Y* and not between the *symptomatic lever W*, which acts on the symptom, $E(Y)$. Nevertheless, many persons, through ignorance or obstinacy, favor symptomatic control, thereby often aggravating the initial control problem. This is such a common error that it represents an obstacle in many organizations and in many control contexts, so much so that Peter Senge considers it a true *systems* ARCHETYPE, which he calls "Shifting the burden," which will be illustrated in Sect. 5.6.

5.3   Effectiveness and Efficiency of Control Systems

The general *conclusion* presented at the end of Sect. 5.1 leads to another question: is there some criterion for assessing the *effectiveness* of a control system; that is, its capacity to achieve its objectives or to respect its constraints, with the maximum precision and the minimum risk of failure? And is there also some criterion for testing the *efficiency* of a control system; that is, the capacity of the system to be effective by utilizing the *minimum number of control levers* activated by processes requiring the *minimum use of resources*?

To answer these questions it is especially important to observe that control systems have their own *utility* and that only the governor-manager can determine the overall *utility* of the system by evaluating its effectiveness and efficiency. To assess the *utility* of the control system it is natural to consider the *effectiveness* requirement as having priority over efficiency, since an ineffective system does not allow the variable Y to achieve its objective; thus, *effectiveness* must be viewed as a *vital feature*, often to be achieved "at all costs." Only in the case of the design of alternative control systems—all capable of achieving the same objectives, with the same risk of failure—must the criterion of efficiency have equal importance to effectiveness in the evaluation. Nevertheless, we must consider that a control system is effective not only if properly sized—including all the relevant variables that influence the objectives—but also if the *levers* have a *range of variation* suitable for the admissible variations of the variables to control.

We must remember a *key principle of cybernetics*, which all evaluations of the effectiveness of control systems must adhere to: the *law of the necessary variety*, formulated by Ross Ashby (1957), according to which the "variety" of admissible states of the Y variables of a control system must be equal to or greater than the "variety" of the reality (disturbances, $E(Y)$ variables), which can alter the values of the variables to control, X, to such an extent as to not permit the objectives to be

achieved. From this it follows that an effective control system must possess control levers, *X*, in *number*, *variety*, and *potency* so as to allow the error to be eliminated, despite the vastness of the environmental "variety." "[Only] *variety can destroy variety*" (Ashby 1957, p. 207). Ashby provides a formal demonstration, but I think it is sufficient to remind the reader how Ashby applies the law of requisite variety for the control of living systems:

> ... the law of Requisite Variety enables us to apply a measure to regulation. Let us go back and reconsider what is meant, essentially, by "regulation." There is first a set of disturbances D, that start in the world outside the organism, often far from it, and that threaten, if the regulator R does nothing, to drive the essential variables E outside their proper range of values. The values of E correspond to the "outcomes" [of the regulator]. Of all these E-values only a few (η) are compatible with the organism's life, or are unobjectionable, so that the regulator R, to be successful, must take its value in a way so related to that of D that the outcome is, if possible, always within the acceptable set ... i.e., within physiological limits (Ashby 1957, p. 209).

In any case, the larger the variety of actions available (*X*) to a control system, the larger the variety of disturbances it is able to compensate for.

> Ashby's law is perhaps the most famous (some would say the only successful) principle of cybernetics recognized by the whole Cybernetics and Systems Science community. The Law has many forms, but it is very simple and commonsensical: a model system or controller can only model or control something to the extent that it has sufficient internal variety to represent it. For example, in order to make a choice between two alternatives, the controller must be able to represent at least two possibilities, and thus one distinction. From an alternative perspective, the quantity of variety that the model system or controller possesses provides an upper bound for the quantity of variety that can be controlled or modelled (Heylighen and Joslyn 2001).

Following Fig. 2.10, if [*X*] stands for the vector of "active" variables and [*Y*(*X*, *D*)] the "passive" ones, which suffer the "variety" of external disturbance variables [*D*], the *control domain* is defined as the admissible range of the variables to control:

$$\left[Y(X, D)_{min} \leq Y \leq Y(X, D)_{max} \right]$$

This domain is obtained as functions of the values of the levers [*X*] and the disturbances [*D*] (to simplify the symbols, the variables in square brackets signify vectors of variables). The law of Requisite Variety enables us to specify the *General Criterion of Effectiveness*: a control system is *effective* in controlling [*Y*(*X*, *D*)] if it includes a number and quality of control levers, [*X*], capable of producing, overall, a sufficiently large *control domain* to achieve the objectives [*Y**]; that is, to eliminate the errors [*Y** − *Y*(*X*, *D*)], producing values such that:

$$\left[Y(X, D)_{min} \leq Y^* \leq Y(X, D)_{max} \right]. \tag{5.1}$$

If the *control domain* is always [*Y** > *Y*(*X*, *D*)_{max}], then the control system is *undersized*: if it is always [*Y** < *Y*(*X*, *D*)_{min}], the system is *oversized*; in either case it is ineffective.

The *General Criterion of Effectiveness* produces the following actions to restore effectiveness:

1. Every *undersized* control system can always be adequately sized by means of two actions capable of raising $Y(X, D)_{max}$:

   (a) Inserting new control levers in [X] and/or strengthening the range of admissible variation of [X] by acting on $g(Y/X)$ and $h(X/Y)$;
   (b) and/or controlling the disturbance variables [D], which will result in a rise in the maximum values of [Y];

2. Every *oversized* control system can always be adequately sized by means of two mirror-image actions that can reduce $Y(X, D)_{min}$:

   (a) Reducing the number of control levers in [X] (deactivation) or weakening some of them by reducing the range of admissible variation;
   (b) and/or controlling the disturbance variables [D] to reduce the minimum values of [Y].

Control actions 1(a) and 2(a) can be defined as *structural* or *internal* because they impact on effectiveness by influencing the internal structure of the control system. Actions 1(b) and 2(b) act instead on the external environment, which represents the source of the disturbances [D]; therefore, I shall call these *environmental* or *external* control actions.

Structural actions depend on the nature of the control system and the available technology, thus on design potential and managerial quality. *Environmental* actions depend, on the other hand, on the type of disturbance that alters the dynamics of [Y]; they can also be carried out by additional autonomous control systems that combine with the main control system in a complementary way.

We must now reflect on which form of control action—structural or environmental—to carry out. I must first note that the general criterion of *Effectiveness* allows us to state the following *General Criterion of Efficiency*: in general, control systems are efficient to the extent they are designed using the minimum number of control levers, [X], necessary to produce an adequate control domain for the objectives [Y*], both in terms of *precision* and the *cost* of control, thereby reducing as much as possible the *external control* of the disturbance variables D.

Secondly, it is appropriate to point out that a number of jointly controlled objectives and a number of control levers simultaneously considered depend on how the *designer* and the *manager* of the control observe reality; the more they zoom in "on a narrower perspective," the fewer control levers there are. The more they zoom out "towards a broader perspective" of the observed environment, the more the control will involve multiple objectives and imply multiple levers.

The eye system taken by itself (zooming in on a narrower perspective) can control the quantity of light that reaches the retina through the contraction of the pupil (first lever) and the closing of the eyelids (second lever); if we also consider the individual, we can add a third control lever: dark eyeglass lenses. If we zoom in to observe an individual in a room we can implement the control with *environmental actions* as well, such as the opening or closing of the shutters and the switching on or off of the lamp. If we consider that the individual must also have the sharpness of vision
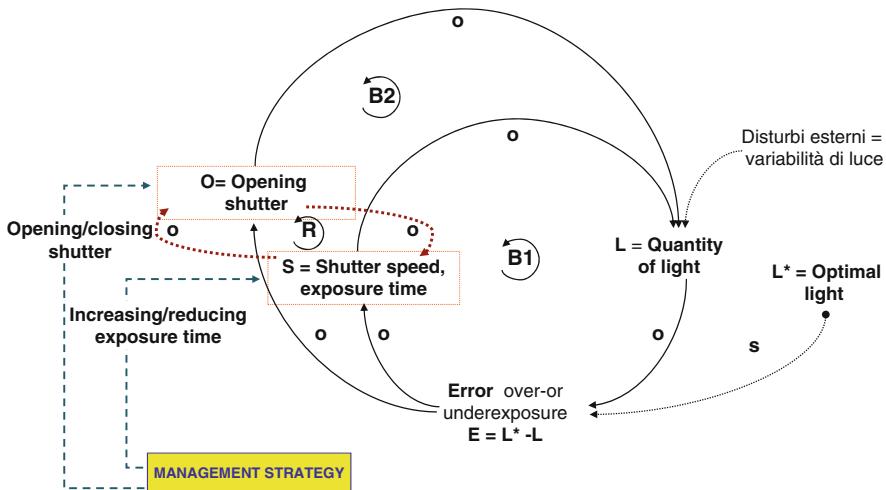
Fig. 5.1 Model of a control system for the exposure in a manual reflex camera

necessary to read tiny characters, then we cannot limit the control to only the quantity of light but must also include contrast as a further control objective. This control system is entirely similar to that for the control of the exposure in manual reflex cameras, which is shown in Fig. 5.1.

Generalizing, the more we zoom out (a broader perspective) the more necessary or useful it becomes to increase the *number* of [Y] variables that the control system must consider simultaneously. However, we must at the same time consider an increasingly higher *number* of disturbance variables [D], with the inevitable result of an increase in the *number* of control levers [X] needed to achieve the objectives. It follows that the larger the environment in which the control takes place, the more numerous and complex are the *structural* and *environmental* control actions necessary to make the system *effective*.

We thus have an initial answer to the questions we started with: the size of the control does not depend so much on the "nature" of the situation we are examining as it does on the extension of the *boundary* that, in delimiting this reality, defines the control objectives the observer can pursue. In other words, every control system is not absolute and objective but *relative* and *subjective*.

5.4   Strengtheners, Turbos, and Multi-levers

It is important here to introduce another important point: when *observing* or *designing* control systems it is necessary to carefully analyze the *nature* of the control levers [X] in order to see whether or not they can vary continuously and for discrete values, and especially if they have a *capacity*, which derives from a *section* and a *speed of transition*. This necessitates the determination of values for the

system's rates of action, "$g(Y/[X])$," reaction, "$h([X]/Y)$," and the characteristics of symmetry, ($h = 1/g$). Only knowledge of these structural elements allows us to study the system's behavior and thus make calculations using the recursive model that produces it or to simulate its dynamics by using some specific tool.

In order to further consider control mechanisms, let us return to the control system for a *bathtub controlled by two independent faucets*, similar to the one in Fig. 4.1. We assume that the system has a positive lever, X2, that increases the water level at a *section* $sX2 = 2$ and a *velocity* of $vX2 = 2$; assuming also that the system is perfectly linear and symmetrical, this means that the *capacity* is $gX2 = 4$ units ($s \times v$) and the water level rises by that amount for each period of measurement.

Let us consider an operator–manager who, finding the tub empty, sets the objective of reaching a level of $Y^* = 500$ units in a given period, $T = (0–50)$. We can easily verify that the objective cannot be achieved *in the time available* since, with a capacity of $gX2 = 4$ units for each "t," the level would only rise by $Y_{50} = 200$ units after 50 instants. The system is *undersized* in terms of the objective; or, from a different perspective, the objective is *oversized* in relation to the available effectiveness of the control system. In this specific example, in order to strengthen the control system we can carry out several *actions* on the system's *effector* mechanism, which we define as *structural actions*; for instance (simplifying the symbols) we can:

1. Increase the *velocity* of the outflow; if the velocity is increased from $v = 2$ to $v' = 5$ per unit of section, with a section lever equal to $s = 2$ there would be an inflow of $g = 10$ units per instant; the objective could be reached in 50 periods. This action can occur through a *structural* intervention that installs a pump to increase the water *pressure* and thus the *velocity* of inflow as well.
2. Increase the *section* of the inflow lever while maintaining the same flow velocity; this adjustment assumes that it is possible to restructure the physical system by increasing the *section* of the water pipes.
3. Add a *new inflow lever* capable of supplying additional water; this implies a more detailed restructuring of the system, which then would have two independent levers for the water inflow in addition to a dependent lever for the outflow.

Environmental actions can also be taken that do not concern the *structure* but the external *environment*. These can again involve the *velocity* of the water inflow by means of a lever or a "device" that is not part of the system; for example, a pump placed in the tank where the water comes from through the inflow lever or a process that brings the objectives to a more manageable level for the system.

Mutatis mutandis, the situation is similar for the outflow process which, if undersized, can be regulated through *structural* and *environmental* interventions.

In general, we can define *boosting* as any *structural* intervention that increases the *capacity* of the lever $g(Y/X)$. *Boosting* can occur through two *internal* interventions, which are individual or combined:
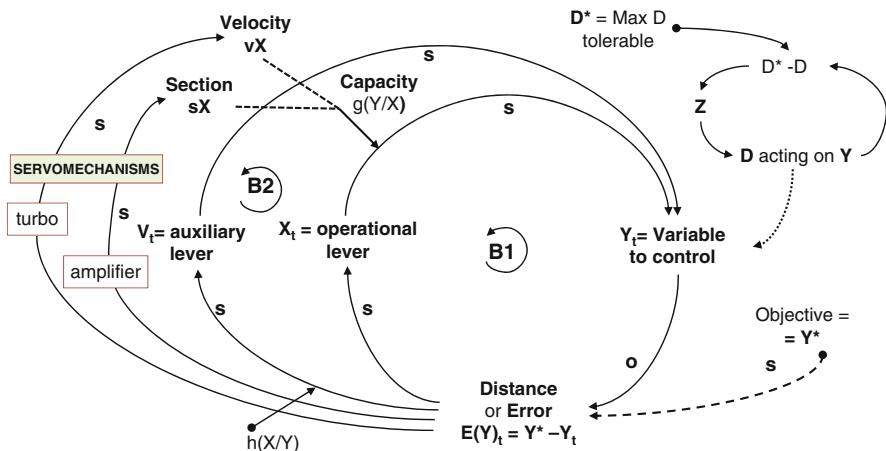
Fig. 5.2 Single-lever control system with booster adjustments

– An *amplifying intervention*, if it acts on $g(Y/X)$ to produce an increase in the section, "*s*," of the lever (the meaning of which is to be specified for any concrete system).
– An accelerator or *turbo intervention*, if it acts to increase the flow velocity, "*v*."

Amplifying and turbo processes can depend on both the controlled variable X and the error $E(Y)$; in this case the control system frequently becomes multilayered.

The *apparatuses* that produce these effects are internal *servomechanisms* that can produce linear dynamics, if their effects do not depend on the amount of error, or nonlinear dynamics, if in some way they are a function of the $E(Y)$ or its components, the level of Y and/or the objective. Nonlinear boosting processes are more familiar to us from our experiences since, as we know, a turbo has stronger effects when there is a large variance with respect to the objective; the effects are gradually reduced when the dynamics of X approach the objective.

A *new control lever*, V_t (itself dependent on the error, $E(Y)_t$), which produces a further flow with respect to the lever X_t, can be defined as an *auxiliary lever* that is *internal* to the structure; if, instead, V_t is not regulated by $E(Y)$, then the *servomechanism* is *external* or environmental.

Applying these internal and external actions to the general model of a control system in Fig. 2.10—assuming only one lever, X, to make the graph simpler—we get the *undersized* control system model, with the possibility of *booster* adjustments, as shown in Fig. 5.2, which, although it seems complicated, includes all the booster elements mentioned above. The initial system is represented by the *loop* [**B1**], to which is added the second lever V_t, which starts the loop [**B2**]; this lever represents an internal *auxiliary lever* which can booster the action of X on the variable Y.

The *amplifier* and the *turbo* are represented instead as links between the error and the action parameter, $g(Y/X)$, which is divided into its two elements: *section* and *velocity*. The external disturbance, D, which disrupts the value of the variable, Y, is

kept under control by an *environmental* control system that, by means of the external lever "Z," carries out an external control process to reduce the undersizing of the initial system. By definition, the *environmental* control is *external* to the main control system. Lever "Z" could also be viewed as a second- or third-order *structural* lever (Fig. 4.13), and thus placed inside the main control system, which then becomes a *multilayer* control system (Sect. 4.5).

Let us not, however, get too far ahead of ourselves. Today all cars have the *turbo*, the well-known servomechanism that increases the pressure of the mixture that enters the cylinders to increase their power in order to increase the car's velocity through pressure on the gas pedal. An external *amplification* increases the car's piston displacement; but given equal piston displacements, cars with a turbo with more thrust are more desirable. A bicycle's amplification mechanism is its derailleur, which allows the gears to be changed so that, with the same pedaling frequency, the speed increases. The turbo effect can only come from the outside, by increasing the muscular strength of the cyclist's legs. An auxiliary lever could be represented by an electric motor that would make the bicycle go faster. The amplification effect is achieved in a sailboat by varying the number and surface area of the sails; the turbo effect cannot be achieved structurally, but depends instead on the wind velocity. An auxiliary lever could be represented by a motor that added its effect to that of the sails.

5.5   Risk of Failure of the Control Process Due to Structural Causes

After having presented some guidelines for observing and designing control systems (Sect. 5.1) we need to deal with the other side of the coin: providing indications for recognizing the many *risks of failure* of the control process. We must begin by considering several errors that very frequently lead to an inadequate control system that risks failing as soon as it starts up. The most frequent of these concern the definition of the variable to control [Y], the control levers [X], and the objectives [Y*]; in particular, the inability to correctly *identify* or *design* the link between the variables [Y] and the levers [X] that control the former in order to achieve the chosen objective. This problem is evident when we consider, for example, the biological control system for thirst that we examined in Sect. 4.2.

Remaining on the topic of objectives, it is useful to consider a risk factor that is often underestimated but that can cause catastrophic damage to the system due to the inability to connect levers to objectives when the levers produce cumulative or contrasting effects. Many problems in piloting an aircraft result from the beginner pilot's unawareness that it is not possible to increase altitude by only using the lever represented by the horizontal tail rudder without at the same time using the throttle lever. When the plane "raises its nose," its speed is reduced; if the engine power is not increased the plane risks stalling. Similarly it is useless to try to treat liver disease by using pharmacological levers that damage the kidneys; try to lose weight

with intensive sports activities without considering the effect of this on your heart; or try to increase national output through levers that tend to increase savings, since postponing consumer purchases results in lower and postponed demand, with the opposite effect on the desired objective. A lack of awareness of and respect for the interrelations between the objectives and the related control levers is one of the most serious types of failure of any control process.

A second factor that can lead to the ineffectiveness of the system in carrying out the objective and achieving control is linked to the insufficient *potency* of the control levers; that is, the incapacity of these to influence Y, taking into account the nature and potency of the *effector* apparatus. As observed in Sect. 5.4, if actions are not taken to increase the potency of the levers that are undersized, the control system will inevitably fail to achieve the objective.

If a river regularly overflows we can control the flood waters by building fixed embankments; this lever only has a *section*, in this case represented by the height of the embankment. It is not possible to also determine a *velocity* for this lever. Even if we build an outlet canal for the water, which activates when the water level rises, we still have a lever with only a *section*. However, if we use a water-scooping system that is activated when the water level reaches the flood level, then we have a control lever whose *capacity* or *potency* derives from both the *section* (number of water-scoopers, diameter of the tubes) and the *velocity*, that is, the pumping capacity. Examples of failures in the controls due to insufficient *capacity* of the control levers are so evident that we need not mention them; we all know it is impossible for a ship to leave the port if the tugboat is not powerful enough, just as it is impossible to quench your thirst on a torrid day with a drop of water per minute. It is nearly impossible for a goal-keeper to stop all the penalty shots; his speed and arm length are too limited with respect to the speed of the ball, and he can only stop the shot when it is too central or slow.

Linked to insufficient *potency* of the levers is another risk factor that can lead to the failure of a control: the incapacity or impossibility of *observing* or *designing* an adequate number of control levers and of ensuring a quality suitable for the variable to control. How can we control the continual rising temperature on our planet (Sect. 6.6 below) How can we stem the gradual desertification around the world? How can we eliminate the periodic invasions of locusts? How do we control the reproductive rhythm of cancerous cells before there is an explosive and lethal increase in them? Based on our current knowledge, these all seem to be uncontrollable systems, even if the efforts to create adequate *levers* for these situations are well known. Some of these levers have been discovered (pollution, forest fires, the spread of algae), some have been developed conceptually (radiotherapy, receptor inhibitors in cells), and others are being experimented (new food plants derived from genetic engineering, education of the population, food and medical assistance). One thing, however, is clear: the variables to control will always be out of control until the basic *levers* that affect these dynamics are not completely included in an effective control system.

Last but not least, we shall consider an important structural factor almost certain to produce failure in the control, even when this is well designed. Section 2.9 showed

that as the *chain of control* is made up of three distinct apparatuses—*effector*, *detector*, and *regulator*—it is necessary to have a *transmission network* for information, however constructed, which connects the three apparatuses and must be considered complementary to the three "machines" making up the *true* control system. Nevertheless, clearly this network must be considered, in all regards, as a specific *autonomous* apparatus, on whose correct functioning depends that of the entire control system. The concepts of *network* and *transmission channels* for information not only concerns mechanical, electrical, or electronic connections among the apparatuses in the chain of control but must also be viewed in a broader sense in relation to the nature of control systems. Here are some examples.

Some hold that the battle of Waterloo was probably lost by Napoleon because he could not activate an important control lever: the positioning of General Emmanuel de Grouchy's 30,000 troops, which represented the right wing of the French alignment. The reason for this was because it was not possible to communicate orders to them and thus carry out the strategy; the channel of communication (the dispatch riders) between the regulator (Napoleon) and the effector of the lever (de Grouchy's troops on the right wing) did not function perfectly. Moreover, it seems that Napoleon advised de Grouchy to listen for the sound of the cannons to understand in which direction to march to bring support; however, de Grouchy stuck to the previous orders to pursue the Prussians rather than maneuver in aid of Napoleon.

A cataclysm that physically separates two populations, prey and predator, would represent a malfunctioning of communications between the effector and detector, impeding the biological interactions among individuals and upsetting the biological control system. No matter how relevant the gap is between actual sales data and the budgeted data, no measure (regulation) could be taken with regard to the sales force (effector) if the head of periodic sales reporting (detector) did not transmit the noted variances to the marketing director (regulator) in a timely fashion. No control of the punctuality of urban buses could be carried out if the reports by the route inspectors (detectors) did not arrive at the municipal transport command center (regulator).

A final striking example reveals the insidious nature of human error regarding information *transmission channels*. As we know, the most accredited theory about the sinking of the Titanic in April 1912, which caused more than 1,500 victims, is that the transatlantic liner collided with an iceberg due to a defect in the rudder (effector), which was not adequately sized and equipped for the rapid evasive action required by the late sighting of the obstacle (Error detection). Recently, however, another fact has emerged. The tragedy was not caused by the delay in detecting the iceberg or by the inadequately sized rudder, but probably was due to a simple error in communicating the control decisions for the evasive action. We read from the diary (only now made public) of the second officer, Charles Lightoller, the ship's fourth most senior officer (who survived the sinking):

> The error on the ship's maiden voyage between Southampton and New York in 1912 happened because at the time seagoing was undergoing enormous upheaval because of the conversion from sail to steam ships. The change meant there were two different steering systems and different commands attached to them. Some of the crew on the Titanic were used to the archaic Tiller Orders associated with sailing ships and some to the more modern Rudder Orders. Crucially, the two steering systems were the complete opposite of one another.

> So a command to turn "hard a starboard" meant turn the wheel right under the Tiller system and left under the Rudder. When First Officer William Murdoch spotted the iceberg two miles away, his "hard a-starboard" order was misinterpreted by the Quartermaster Robert Hitchins. He turned the ship right instead of left and, even though he was almost immediately told to correct it, it was too late and the side of the starboard bow was ripped out by the iceberg. The steersman panicked and the real reason why Titanic hit the iceberg, which has never come to light before, is because he turned the wheel the wrong way," said Lady Patten who is the wife of former Tory Education minister, Lord (John) Patten (Richard Alleyne 2010, online).

The quote speaks for itself.

There are many other risk factors of control; however, I wish to conclude this section by pointing out another element that often makes control problematic. Though taken up at the end, it is certainly not last in terms of its importance and the disastrous effects it can have on the control process.

We saw in Sect. 2.9 that an essential component of the chain of control is the process of *detecting error*—that is, the deviation $E(Y)$—which is a process set off by a *sensor* apparatus that is often very complex. We need to realize that without an efficient detection process for error, or when the process is unreliable and imprecise and operates with a non-admissible tolerance, it is not possible to calibrate the regulation of X so that it acts appropriately on Y; as a result control is not possible or is extremely imprecise. This risk emerges not only in cases of possible malfunctioning of the mechanical or electrical instruments for error detection but also, and above all, when the detection is carried out by the operator–manager using his own senses and knowledge.

There is no need here to use overly sophisticated examples such as the lack of control of solar batteries on the space probe due to the malfunctioning of the charge detection; we need only point out all the times our car tires have gone soft because, being distracted, we had failed to check the pressure; what a relief today to have an automatic pressure gauge that signals any problem, thus allowing us to promptly carry out the necessary control.

5.6 Risks of Failure of the Control Process Due to Improper Levers: "Shifting the Burden" ARCHETYPE

Much more subtle are the risks from choosing the *wrong lever*, which is not able to control the variable to be kept under control. There are several variations of this situation. The most evident is choosing levers that only act on the *symptoms*. The action of symptomatic levers, X, starts the control positively; but since Y is not directly controlled, the error is not eliminated but instead returns even stronger later on. This unsuccessful control takes on two forms that are so frequent they are recognized as widespread recurring *archetypes* in every context (Sect. 1.6.7). The first variant of a failure due to the erroneous identification of the control levers can be represented by two archetypes similar in their effects: the ARCHETYPE of the "Fixes that Fail," which operates in all contexts where control favors short-term levers, and the more general ARCHETYPE of "Shifting the Burden," which acts in
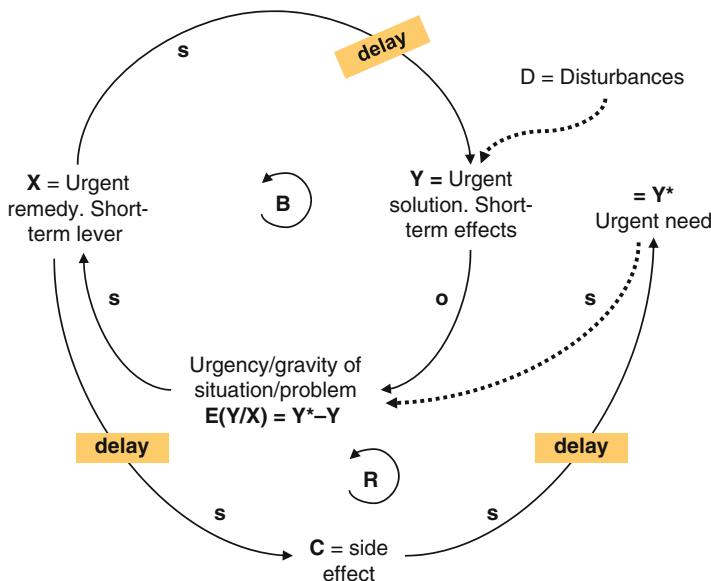
Fig. 5.3 "Fixes that Fail" ARCHETYPE applied to the control process

situations where control is based on symptomatic levers (this can be viewed as a generalization of the previous archetype).

The "Fixes that Fail" ARCHETYPE is applied to control systems that can have several levers. Some of these are symptomatic levers, which act immediately on Y or directly on $E(Y)$, though their effect can be short term so that, once eliminated, the deviation returns after a short period. Other *levers* are instead *structural* and long term and often act with a delay; however, their effect is permanent. Considering the period needed for them to act, the former can be called *contingent* levers, while the latter can also be called *strategic* levers. Figure 5.3 describes the archetype of the "Fixes that Fail."

Pollution has to be reduced? The simplest *symptomatic* lever is to block cars from circulating. This lever has an immediate effect on pollution, as stopping cars from circulating reduces dangerous emissions. However, clearly this affect is *short term* because the economics of a region is based on the rapid displacement of people, and thus cars cannot be blocked from circulating for a long period of time. Quite different is the *structural* lever that rationalizes traffic by redesigning the road system to avoid traffic jams; or the lever, even longer term, whereby the subway or rail system is improved. An even stronger effect would be produced by the *lever* represented by the spread of electric automobiles. These are long-term levers not only on account of the time needed to produce the necessary infrastructures but also because of the need to modify citizens' attitudes toward using their cars.

Drug consumption must be controlled? The shortest-term *symptomatic* lever is the usual one: try to arrest drug traffickers and drug dealers; a similar measure

would produce a short-term improvement since, as we observe in real life, the spread of drugs almost always immediately recurs in other ways. This depends not so much on the possibility of business activity in this area but on demand. A *structural* lever with permanent effects, but which is applicable only in the long term, would be to improve education at all levels about the harmful effects of drugs, accompanied by social conditions that provide more positive incentives.

When a control system is structured only with contingent, short-term levers the control can fail, because in the long-term unforeseen consequences always arise that manifest themselves as disturbances that do not permit the objective to be achieved and stably maintained, thereby generating new, more serious deviations that make a continual readjustment of the solution necessary. The "Shifting the Burden" ARCHETYPE shows the recurring risk of control failure when the control system is designed solely to eliminate the symptom, so that the user–manager, perceiving a temporary improvement in the unpleasant sensations linked to the symptom, neglects to activate the levers that would lead to a long-lasting control of Y, thereby causing a net worsening in the system's performance, with often dramatic consequences (Kim 1999). If we try to control a fever with ice on our forehead without treating the infection with antibiotics, the system will improve for a few hours but the infection will spread. If we try to resupply a warehouse with extraordinary production, without re-planning production and sales, the rundown of stocks will inevitably resurface.

Figure 5.4 shows the structure of the "Shifting the Burden" ARCHETYPE applied to the control process. When the control system is structured to reduce the intensity of $E(Y)$, but the structural levers have a long operating time or involve delays, then $E(Y)$ appears as a serious symptom to be eliminated through symptomatic levers, by forming loop [**B2**], which can cause a delay in the activation of the strategic levers for the purpose of the structural control [**B1**]; this delay can lead to collateral effects that further hinder the search for strategic levers, thereby producing a new and more serious variance through the strengthening action [**R**]; systems thinking—precisely because it brings out the perverse *modus operandi* of this archetype—favors and accelerates the search for strategic levers for an effective control. Fortunately this archetype does not always act in all control contexts, its action is more easily encountered the more intense the unpleasant effect associated with the symptom is, so that in order to eliminate it we end up by reducing the efforts to activate structural levers, with the consequent delay in achieving the objectives.

My father has a bad back because he weights too much, and the weight compresses the vertebrae in the spinal column. The pain he feels, E^1, comes from the excessive weight (Y) compared to the weight the spinal column can support (Y^*). The structural solution would be for my father to reduce the amount of food (X) he eats through an appropriate diet in order to reduce his weight (Y), since obviously the excess weight impedes him from moving freely. However, the effects of the diet are seen only after a long delay, and thus the excess weight ($Y > Y^*$) does not diminish fast enough. The elimination of back pain thus becomes the overriding variable, E^2, which requires that E^1 be urgently eliminated; for this
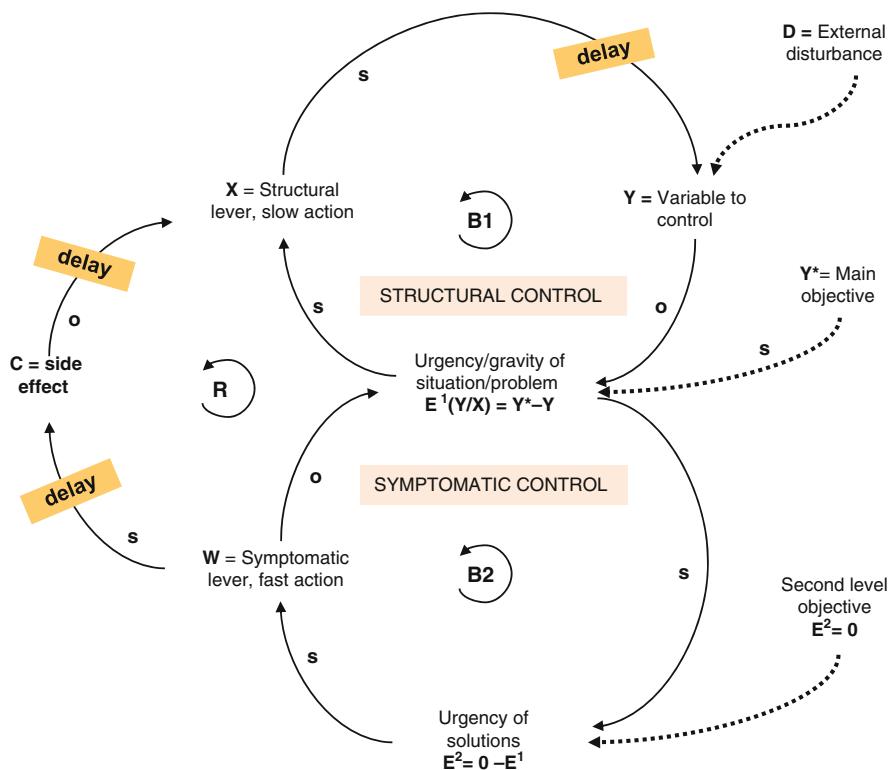
Fig. 5.4 "Shifting the Burden" ARCHETYPE applied to the control process

reason my father turns to a symptomatic lever: pain killers and physiotherapy (the structural lever in the form of a spinal column operation is discarded a priori).

The symptomatic levers act quickly, so that the desire to eat less (X) diminishes, and therefore the pain reappears. The continued use of pain killers produces dangerous side effects on the liver as well as reduces the ability to concentrate, thereby making work more difficult.

The "Shifting the Burden" ARCHETYPE, which generalizes, so to speak, the "Fixes that Fail" ARCHETYPE, has particular importance, since it favors the tendency to prefer short-term symptomatic solutions that immediately produce visible effects, in this way putting off strategic, definitive solutions, thus laying the basis for the production of permanent negative effects.

Therefore, this archetype reveals our natural tendency to prefer present, individual or local advantages that can be immediately enjoyed to future ones we hold as too far off in time, as shown by the "Short-term Preference" ARCHETYPE (Sect. 1.6.6). The "Shifting the Burden" ARCHETYPE is quite common, and it produces its negative effects in quite different situations, individual, social, or organizational.
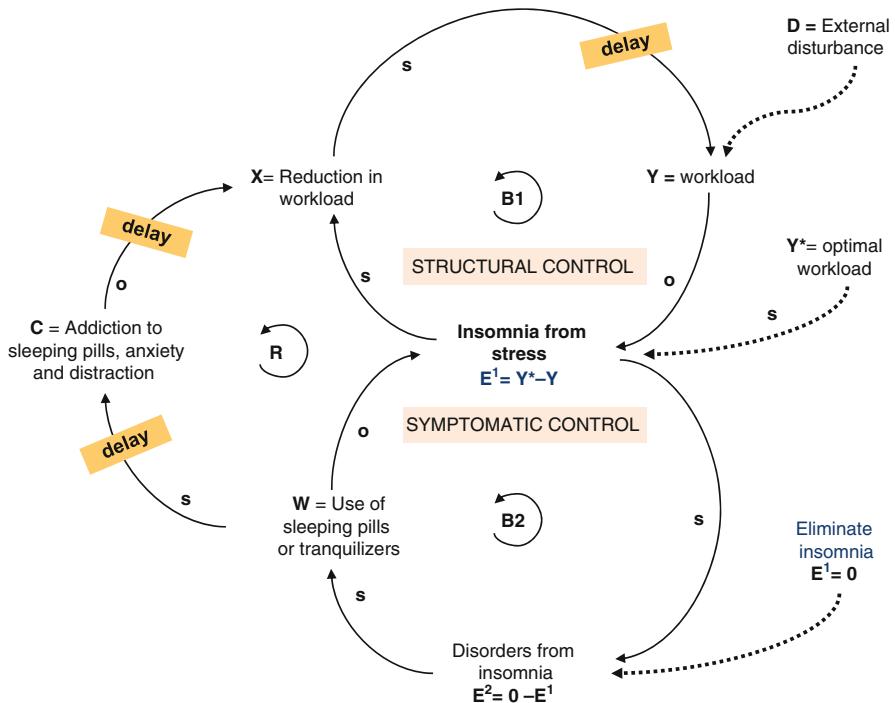
Fig. 5.5 Burnout from stress revisited (see Fig. 1.7)

Figure 1.7 presented the burnout-from-stress problem, which can be translated into the model in Fig. 5.5.

Many growing companies are faced by a strong rise in demand but are not able to provide this through present production. The optimal solution would be to increase the productive capacity through appropriate investments. Since the demand pressure is always higher and there is an increased risk of losing customers, and since the growth in the productive capacity can occur only with a long delay and a substantial investment of financial capital, companies find it much easier to supply themselves externally, by providing other firms with the know-how to produce their products. This problem is so widespread that it can be described as "Shifting the Burden," as shown in the model in Fig. 5.6.

Outsourcing "wins" over increased productive capacity, but it has two negative consequences: the first is an increased dependence on the outside supplier; the second, a reduced urgency to raise capital to finance the increase in productive capacity. These two effects, acting jointly, postpone for a long time the company's productive growth, thereby increasing the risk linked to outside supply.
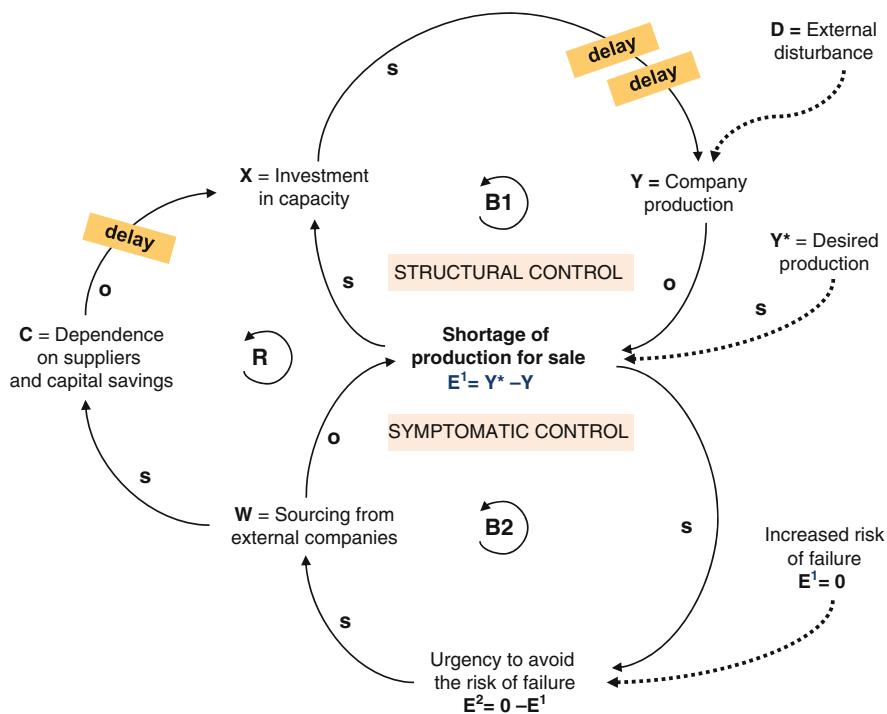
Fig. 5.6 Victory of external sourcing

This archetype inexorably appears in many other situations involving organizations and firms. Figure 5.7 represents the model that explains the delays in organizational growth. In fact, managerial problems are often resolved by recourse to outside consulting which, on the one hand, can solve the problem in the short term, but on the other inevitably provides a symptomatic solution, thereby delaying the true long-term strategic solution, which consists in training managers with ad hoc in-house courses to favor their professional development.

The "shifting the burden" ARCHETYPE has an interesting extension. We know that control systems normally try to lead Y to the values Y^* by means of the levers X; however, in this attempt the elimination of the error becomes difficult when the system undergoes strong disturbances from the external variables, D, which impede the maintenance of equilibrium. Every attempt to readjust the equilibrium, disturbed by D, through the use of the levers X ends up being only temporary, and readjustments are then continually needed.

It is thus natural to ask if the variable Y that the main control system tries to control cannot be considered, in all respects, a *symptomatic variable* controlled by the *symptomatic lever X*; and if, in order to achieve Y^* and maintain a stable equilibrium, it is not necessary to instead intervene on the external disturbances D, which represent the true *strategic variable*, whose control is essential for
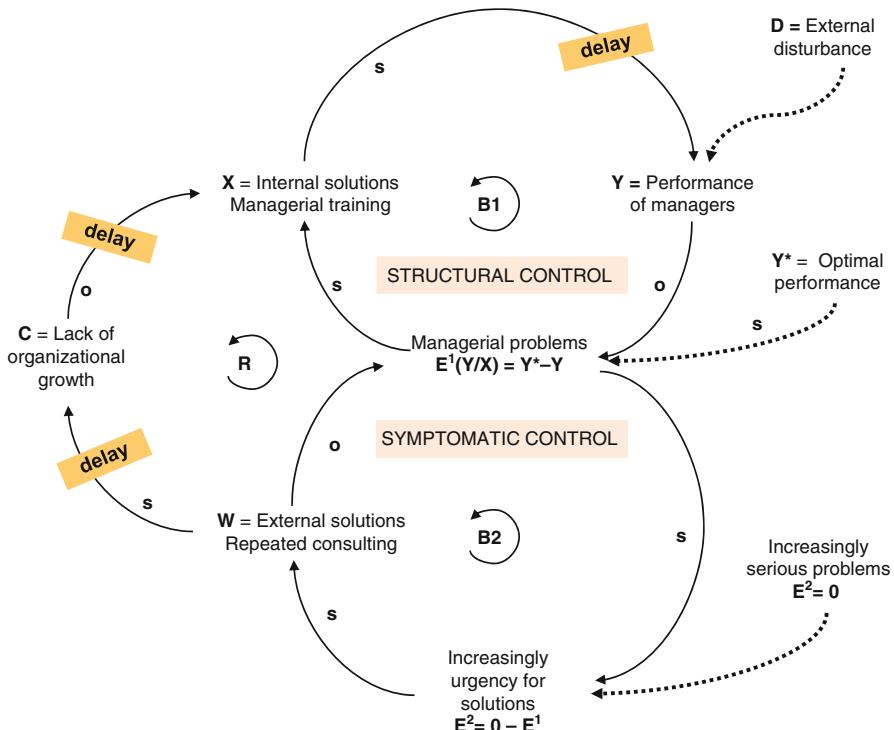
Fig. 5.7 Victory of consultants

achieving the control of Y. Based on what we have seen in Sect. 5.4, this would necessitate the activation of a second external control system on D through the structural levers Z, which are able to act on the disturbance variables, as illustrated in Fig. 5.2. Figure 5.4 would still be appropriate, with the variant that the loop [**B2**] would constitute the long-term structural control system of D, while loop [**B1**] would represent the system that initiates the short-term symptomatic control of Y, as better shown in the model in Fig. 5.8.

The "Shifting the Burden" ARCHETYPE can thus operate in a descending manner, zooming in, when the structural control of Y is replaced by a symptomatic control, as in the model in Fig. 5.4; or it can operate in an ascending manner, zooming out, when the original control on Y is viewed as symptomatic and is supported by the structural control on D, as illustrated in Fig. 5.8. We can also create a *control chain*, where the base-level disturbances are controlled by a higher-level system. Any disturbances at the new level could be controlled at an even higher level, all the way up to the maximum level of admissible structural control, after which no new structural level can be employed.

There are an endless number of examples of this expansion of the original control system's confines. If the control of the level of rice fields is compromised
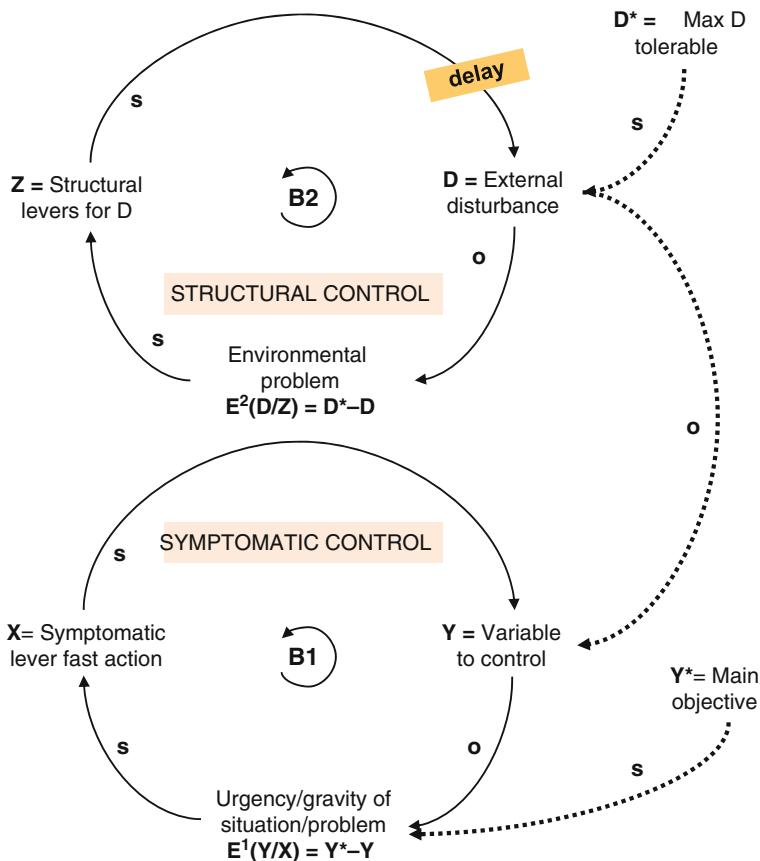
Fig. 5.8 Structural control of external disturbances

by the variability in the river level, it is necessary to construct embankments; if these are not sufficient then the course of the river must be changed or a dam constructed; and if even this is not enough, then the course of another river can be changed to keep the artificial basin adequately supplied with water. If I cannot control the quiet in my home because of noisy neighbors (external disturbance), I can ask them to behave in an educated manner or I can even call in the police. In conclusion, the more the system is expanded, the more the control levers of a given level become symptomatic levers for a higher level of control; the higher-level levers are always more effective, but also slower and more costly, so that it is fundamental to carry out a cost–benefit analysis (Sect. 4.7).

Linked to the risk of failure due to the wrong choice of levers is the wrong specification of the action function $g(Y/X)$ and of the complementary reaction function $h(X/Y)$, taking into account that not all control systems are perfectly symmetrical (Sect. 2.3). In all the examples in the preceding chapters I have chosen to represent "g" and "h" as constants, in order to simplify the simulations and make

them readily understandable. It is intuitively clear that, even with this simplified assumption, the quantification of the rates requires particular attention, as regards both the design and realization of the system and the construction of its model. If we abandon this simplifying assumption and assume instead that $g(Y/X)$ and $h(X/Y)$ are functions of both X and Y, then the control system can also take on a rather complicated mathematical structure, since the relations between the variations in X_t, Y_t, and E_t must be described by differential or difference equations. In any event, we must always remember that without the correct formulation of the functional links among the fundamental variables of the control system, and without a precise quantification of the parameters included in these functions, there is a high risk of failure for the control system.

5.7 Pathologies of Control: Discouragement, Insatiability, Persistence, and Underestimation

I shall conclude this examination of the characteristics of human intervention in control processes by presenting three recurring situations, three ARCHETYPES, which reflect so-called *pathological situations* regarding the use of control systems. The first situation is the "discouragement" that often overcomes the governor–manager when he is not able to achieve the objectives he has set for the system. We know from Sect. 5.4 that control systems are designed and created to achieve the objectives, Y^*, and that the governor–manager first turns to the operational levers; if the objectives cannot be achieved by those alone, he then turns to the auxiliary, extraordinary, and structural levers; and in some cases he even uses symptomatic and short-term levers when the action times of the others are too long. Nevertheless, in many cases where the manager–governor cannot completely control Y and the error is not eliminated, he becomes discouraged and turns to a fallback solution: lowering the level of the objective Y^* so as to eliminate $E(Y)$ without achieving the optimum value for Y.

This failure of the control is very common and can be observed in many different situations in which man manages the control systems. For this reason this situation represents a true systems ARCHETYPE, called "Eroding goals." This archetype is illustrated in the model in Fig. 5.9, where the loop [**B1**] represents the control system that must achieve the objective Y^*; loop [**B2**] represents the process by which the manager–governor decides to achieve less lofty objectives rather than the chosen one, thereby creating an "alibi" for not making the correct regulations and allowing the system to degrade. If "Eroding goals" is allowed to act undisturbed, a "spiral of degradation" inevitably is triggered regarding the global performance of the control system. The erosion of the *goals* or of the *performance standards* usually is accompanied by harmful side effects, which are not always easy to identify because they appear with a delay in the long term. When it is suspected that this archetype is operating, it is thus always necessary to make an effort to identify the side effects, at least the most evident short-term ones. The lower box in
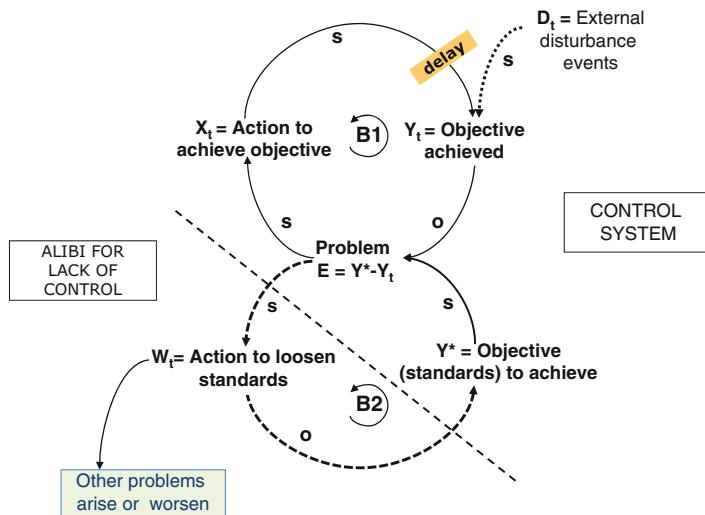
Fig. 5.9 ARCHETYPE of "Eroding Goals"

Fig. 5.9 represents the possibility of side effects, such as the occurrence or worsening of problems. The side effects could also be viewed as external inflows or disturbances; in this way specific control systems could be identified to mitigate these harmful side effects, as illustrated in the diagram in Fig. 5.2.

Nevertheless, despite its clear-cut danger this archetype is found very frequently in decision-making contexts, especially when the manager–governor of the control system is unable to achieve the objectives, judging them to have been too lofty in the first place and being content to accept more modest objectives in the original control system. It is incredible to witness in the papers how this archetype is at play even on a daily basis in families when the behavior of relatives and children is allowed to deteriorate to unsupportable levels of boorishness.

Air pollution has exceeded the tolerance threshold? Then raise the threshold through legislation. The high level of coli bacteria in our waters is endangering bathing? Then raise the tolerance limits with a local decree. Air emissions are raising the temperatures on the planet to dangerous levels because the Kyoto protocol is not enforced? Then let us put off applying it for a few years. Students are too rowdy (they curse, rebel against their teachers, are badly mannered at home)? "They are only kids" … but then we should not be surprised if schools are out of control and bullying has become an ordinary occurrence. Our prisons are overflowing and the conditions of inmates are inhumane? Then we can raise the social tolerance for small thefts, scams, violent episodes, and grant pardons and amnesties. The prisons will empty (the system initially improving) but they will soon become overcrowded again. Some youth have caused damage during the procession (game, film, class outing)? "These are only small fringes" … but let us not be surprised if the areas around stadiums or movie houses become
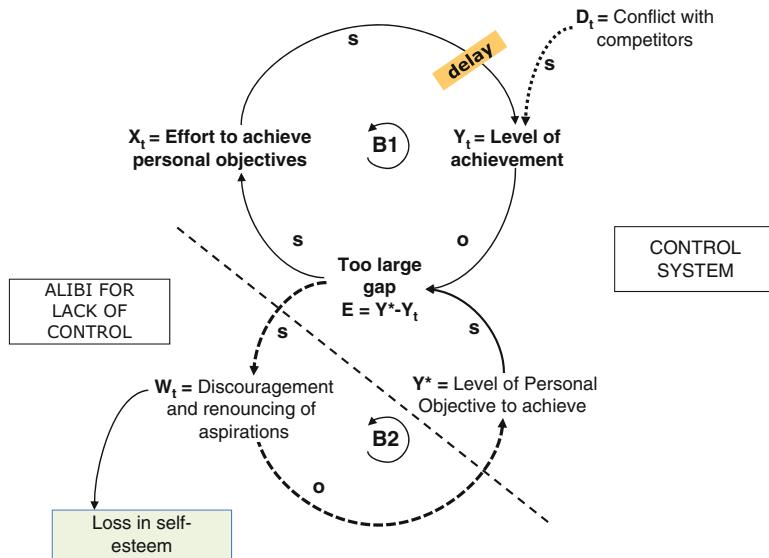
Fig. 5.10 ERODING INDIVIDUAL GOALS and discouragement

ungovernable (for more, see Mella 2012, pp. 240–241). In our daily lives we all fall victim to the temptation to become discouraged when faced with control difficulties. Instead of remaining firm in our aspirations and objectives, we often prefer to sacrifice our ambitions rather than make the effort needed to achieve our objectives regarding sports, study, or work. Figure 5.10 illustrates the tragic case of a decrease in motivation for personal improvement, with the risk of a loss in self-confidence.

The model in Fig. 5.11 shows how the Eroding Goals archetype functions in a common situation that leads to a decline in the quality of services. When the gap between the desired and actual quality is too large, quality control activities are insufficient. Rather than insist in using all the levers available to maintain the original standards, the firm prefers to reduce the acceptable standards. This situation occurs very often in public companies that provide services to customers; however, when this occurs in private companies, which are subject to competition, the decline in the objectives leads to the side effect of a loss of business from unsatisfied clients.

How should the governor–manager behave when he becomes aware of eroding objectives? Here are some possible actions he can take:

(a) First of all, it must be clear that the problem of achieving an objective or a given standard always arises in control systems.
(b) It is then necessary to evaluate whether achievable objectives and realistic performance standards have been set.
(c) A constant focus on the objective and standard must be imposed, and steps must be taken to react to achievement errors, however modest.
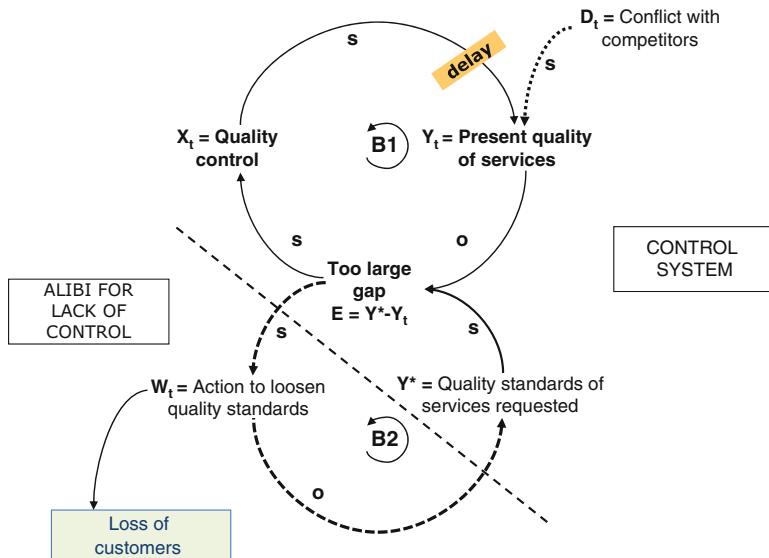
Fig. 5.11 ERODING GOALS and degradation in quality

(d) It is not enough to manipulate the short-term decision-making levers but to work to identify and utilize the long-term levers which, by modifying the generating system, can definitively solve the problem.

It is useful to note that the ARCHETYPE of "Eroding Goals" has a mirror-image variant that we can define as the "Strengthening Goals or Insatiability" ARCHETYPE. This variant is very important, even if in the context of control systems it has not yet received the proper attention. This archetype (illustrated in Fig. 5.12) acts whenever a better result than the objective is achieved and the governor–manager decides not to eliminate the gap by reducing the action of the control lever (thereby producing the loop [**B1**]), but instead to raise the level of standards (loop [**B2**]), thus once again increasing the distance, $E(Y)$, and forcing loop [**B1**] to act more intensively.

Those who are subject to this archetype have no way out. No matter how hard they try they will *never be satisfied*; and even if they do not become frustrated themselves, they end up subjecting their colleagues to an unbearable performance stress. The harmful effects of this archetype are instead felt in organizations whose managers continually raise the performance objectives, as well as in families whose children continually pester their parents for more free time, or whose parents badger their children to get better grades or do better in sports.

Anyone obsessed with exercising an increasingly precise control and who is never satisfied should remember Voltaire's motto: "*the best is the enemy of the good.*" Nevertheless, in many cases "Insatiability" produces positive effects; for example, it characterizes most of the work of scientists, who are never satisfied and
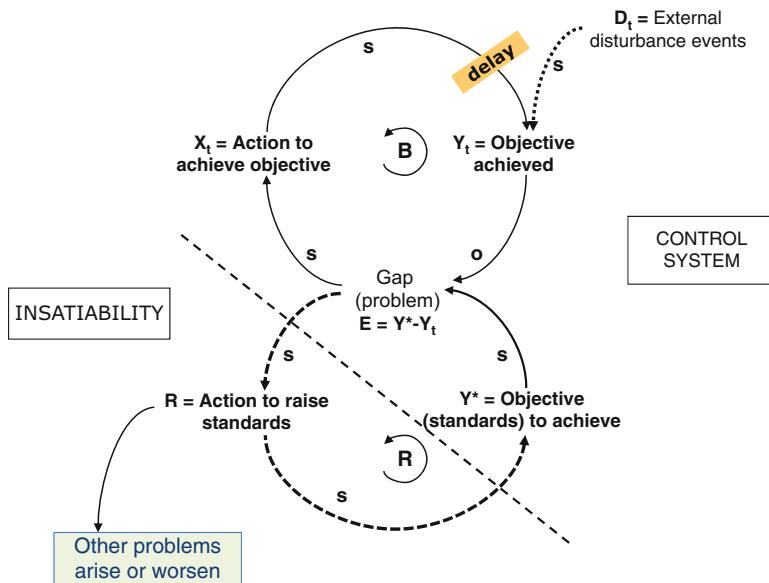
Fig. 5.12 ARCHETYPE of "Strengthening Goals or Insatiability"

fulfilled by their knowledge, of researchers, always thirsting for new discoveries, and of athletes, who are always struggling to set new records.

Figure 5.13 shows how this archetype works for athletes and their trainers, both of whom are searching for continually improving results, which at times harms the athletes themselves, who are subjected to more and more intense training sessions.

The ARCHETYPE of "Insatiability" is often accompanied by another type of equally serious pathological behavior: the *persistence of the controller* with regard to the *controlled subject*. I shall call this the ARCHETYPE of "Persistence," which is shown in Fig. 5.14. It is as equally frequent as the INSATIABILITY archetype and perhaps even more dangerous.

The *pathological* aspect consists in the erroneous perception of the size of the error by the manager, who overestimates the error and brings into play any lever that could possibly eliminate it, demonstrating by this behavior his intolerance toward even the slightest deviation. This behavior represents a true, often *pathological persistence* in the control action. If the objective of the control, *Y**, is the behavior or the performance, *Y*, of some individual, and the *controlled* subject produces even a minimum error, $E(Y) = Y^* - Y$, the manager immediately punishes him, thereby demotivating, frustrating, and, in extreme cases, even harming the controlled subject. Children are the first to be affected by this archetype when parents, teachers, or instructors (governors) say: "Do this!" (objective), and then remain there observing the child (manager) trying to achieve the objective, persisting in punishing him if he cannot "maneuver the levers" to achieve the goal.
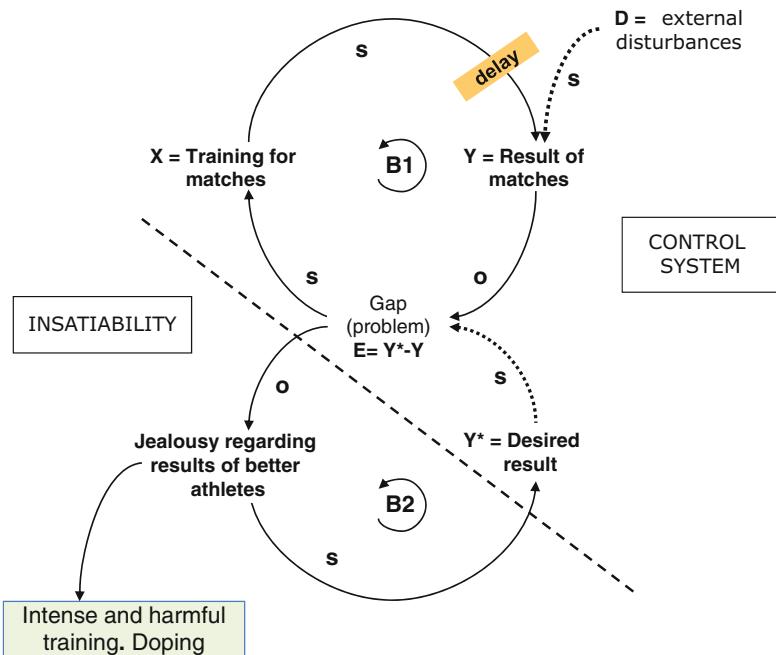
Fig. 5.13 Athletes' INSATIABILITY for victory
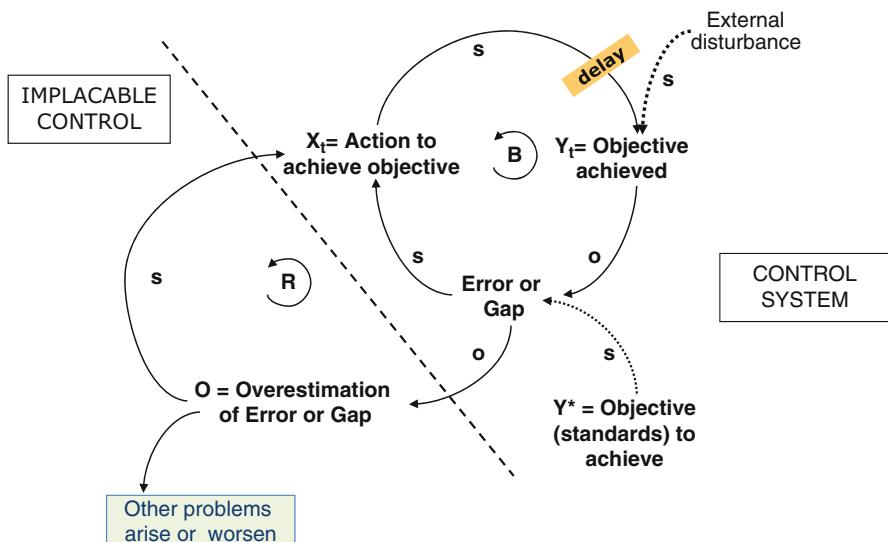
Fig. 5.14 ARCHETYPE of "Persistence"
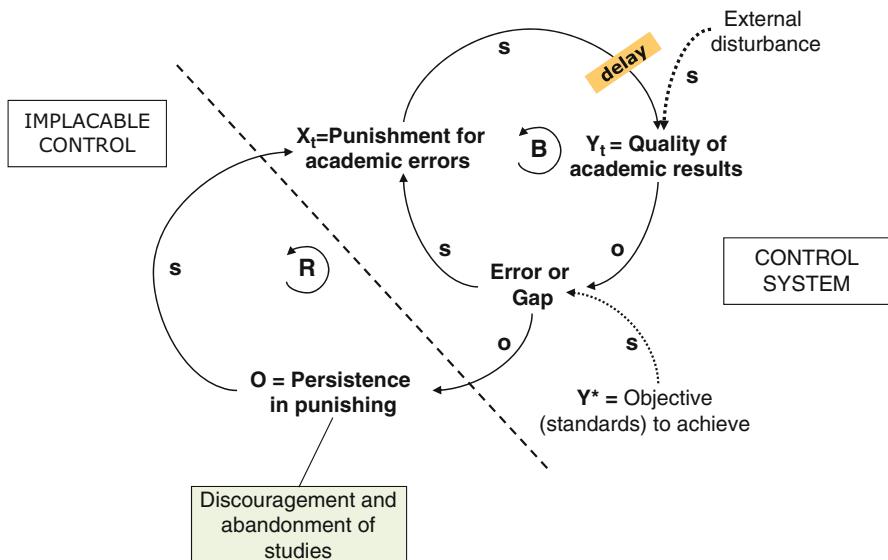
Fig. 5.15 "Persistence" in trying to eliminate academic errors

Examples of how this archetype operates are clear to all of us: the smallest grammatical error is punished with a blue line underscore or a bad mark; the slightest formal mistake in a tax return leads to a stiff penalty; even a minimal excess of speed is sanctioned, even harshly, by the highway police who, perhaps from a hideout, try to catch even the most insignificant infraction. These exaggerated examples are meant to show clearly that this archetype reveals a control pathology whose use should be eliminated.

With reference again to the educational system, Fig. 5.15 illustrates how the archetype of persistence acts in obsessively attempting to eliminate academic errors, punishing students for even the slightest mistake. The students' insatisfaction leads in many cases to students dropping out of school, and even to more serious situations, which often end up in the crime section of newspapers. The two archetypes described above also produce harmful effects in universities. Students know all too well that their inability to pass certain exams is not due to the difficulty of the subject-matter but to the intransigence of the teacher. Who does not remember at least one personal instance of this?

To conclude I shall present a final anomalous, even *pathological* aspect to human intervention in control, which represents the mirror image of the ARCHETYPE of "Persistence." This aspect is equally frequent and perhaps even more dangerous. While "Persistence" leads to an overestimation of the error, leading to the use of punitive levers that are often exaggerated, in many other cases we often observe the manager *underestimating* the amount of error and *acting with a delay* in *regulating* the levers, resulting in the failure to achieve the objective or to achieving it with a delay, thereby causing (even serious) damage.
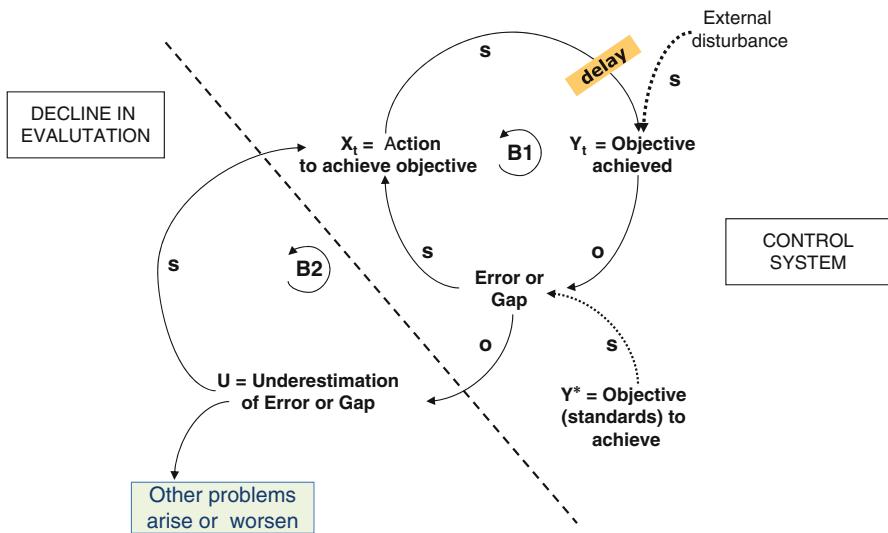
Fig. 5.16 "Degradation of the error Assessment" ARCHETYPE

There is always the same traffic on the road, and if we want to arrive in time at the airport we have to leave the office at 6 p.m.; but there is always someone who is not ready: "We'll get there in time just the same ..." Instead, we arrive when the check-in has already closed. Are you always late for appointments even though you feel you have left in time and there is nothing stopping you from leaving a few minutes earlier? Even though you could be more careful, do you always wait too long to send your purchase order, so that your supplier is forced to deliver it late, each time causing you to stop production?

Such behaviors are not due to bad faith, ignorance, incompetence, unwillingness, or irrationality but to the malfunctioning of the sensor and regulation apparatuses, which are not able to correctly judge the distance, $E(Y)$, between the present position, Y, and the objective, Y^*, and to properly regulate the system's levers. I propose calling this general phenomenon the ARCHETYPE of the "Degradation of the error Assessment," which is illustrated in Fig. 5.16.

This archetype is clearly operating when there is a temporal objective; the erroneous assessment of the "time remaining" leads to delays and to even more serious damage. However, this archetype has a more general action that concerns all types of variables to control. Do these words sound familiar: "Check how much gas we have"; "There's enough still to go a long way." The result: waiting for the tow truck to arrive. "Take the dog out to pee"; "In a moment, there's still time." The result: you have to clean the floor. "It's time to leave"; "In a little while, we'll make it in time." The result: you arrive after the film has started. "Study because your exam is coming up"; "Don't worry! I'll manage without any problem." The result: you fail the exam.
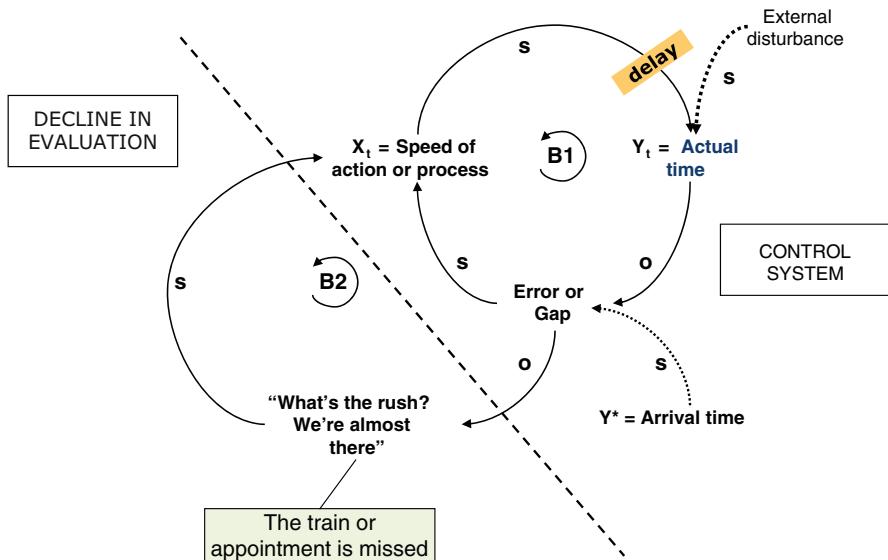
Fig. 5.17 "Degradation" of the error Assessment in the habitual latecomer

I shall conclude this brief presentation of the pathologies of control with the model in Fig. 5.17, which highlights why many people are latecomers "by nature." This is not a question of a lack of volition or of scant respect for timetables; rather it represents the systematic underestimation of the distance between the actual time and the time of the appointment, so that there is delayed or insufficient action on the levers that could accelerate the actions needed to arrive promptly. I would finally point out that persistence as well as the underestimation of the error gives rise to side effects that can aggravate the situation, as is explicitly indicated in the lower part of the various figures showing these archetypes.

5.8 Problem Solving and Control Systems

To broaden our understanding of control systems that every day influence our behavior, it is useful to emphasize the relation between the discipline of control and problem solving, the process by which individuals and organizations recognize and solve their problems (Mayer 1992). We know that "every day and all through the day" we are bombarded by all types of problems to solve in order to live a dignified life (Robertson 2001). Whatever the problem—where to eat and what to wear, not enough money for a project, what name to give our second child, the flat tire to fix on our car, the choice of a new job, a stopped-up sink, etc.—they all have one thing in common: they describe a gap between our present state and the one we have lost or are trying to reach.

Technically speaking, a *problem* is the perception of a gap between the present situation, which we are not satisfied with and would like to improve, and a desired situation, an objective to achieve; or it reveals a worrisome situation we would like to bring back to optimal conditions: "*A problem exists when someone identifies a desired change in a situation*" (Harris 2002, p. 25).

In this sense, the solution to a problem is equivalent to producing a change, and managing solutions to problems can be conceived of as a specific and circumscribed process of change management (Sect. 7.12). In fact, the *solution* of the problem implies the choice of some action to take to try and eliminate the gap between the present and desired situations. The actions we can undertake—called the *courses of action*, action alternatives or, simply, *alternatives*—must be carefully evaluated in order to estimate, with the maximum precision possible, the results, effects, and consequences of each action in terms of the objective, taking into account the possible *states of nature* (that is, the external events)—which do not depend on the decision maker—that can create disturbances in reaching the desired state. The act of deciding is not a simple procedure for choosing an alternative but rather a complex process made up of several phases:

(a) Identifying the existence of a problem and specifying the objectives to pursue to modify the actual state (problem finding):

> Many organizations and their managers drive toward the future while looking through the rear-view-mirror. They manage in relation to events that have already occurred, rather than anticipate and confront the challenge of the future (Morgan 1988, p. 4).

(b) Recognizing the type of problem to solve (problem setting).
(c) Redefining the problem; that is, restructuring the assumptions of the problem and gaining a new perspective on it:

> Problem finding is concerned with identifying the problem. Many problems seemed to be apparent: lots of complaints about pay rates, absenteeism, high turnover of personnel, low standards of efficiency, marked aggression against management, etc.
> Redefining the problem: the boundary-examination technique was used to restructure the assumptions of the problem and produce a new perspective on it. (Proctor 1999, p. 85).

(d) Searching for information (even if it is forecasted information) compatible with the cost of the search and, based on the information, carrying out the subsequent steps.
(e) Searching for technical, legal, and economic constraints to the admissible alternatives, considering the external environment and the internal resources, and delimiting the area of admissible solutions; there is need for *awareness* of the system of *values*, *rules*, and *policies* used in evaluating the identified alternatives.
(f) Choosing evaluation criteria for each alternative based on rules (subjective-intuitive or objective-formal). If a single criterion is chosen, the decision is *single-criterion* or *single-objective*; if there are multiple criteria the decision is *multi-criteria* or *multi-objective* (Sect. 5.10.1).
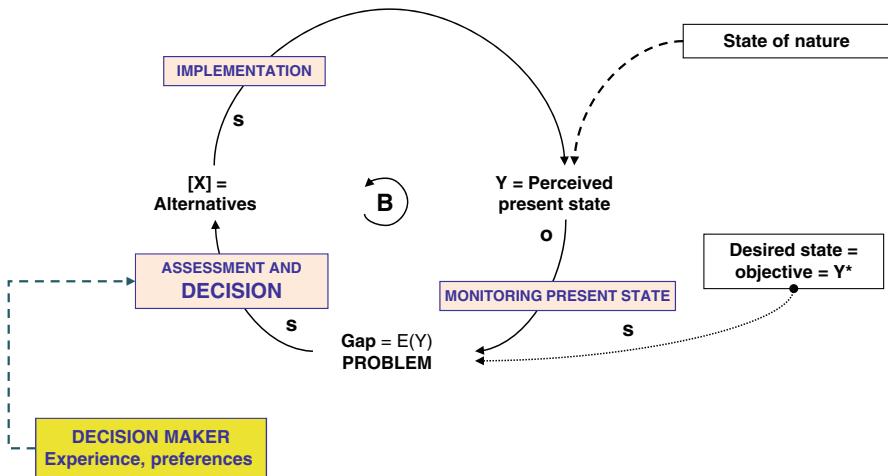
Fig. 5.18   Model of problem solving as a control system

(g) Identifying and evaluating the various alternatives based on the chosen criteria, taking into account the results each alternative permits, based on a given process of forecasting and calculation.

(h) Choosing the best alternative, anticipating the circumstances that most likely could condition the results (non-controllable variables or states of nature); ensuring that the choice is *compatible* with the system of values, policies, and rules shared within the organization.

(i) Implementing the chosen alternative; that is, carrying out a concrete process of action that modifies the present situation in order to achieve the desired one.

Along with implementing the alternatives, the gaps between states and objectives must be controlled; this represents the phase where we evaluate whether the process of change is capable of modifying the present situation in the desired manner.

When the implementation is not capable of attaining the desired results, new problems arise and the process becomes cyclical. The process by which solutions are sought for problems is called problem solving (Shibata 1998). Using the terminology we have just presented, problem solving can be represented as a control system by means of the model in Fig. 5.18.

The reason for this identification is immediately clear if we consider that a problem is nothing other than the gap (deviation, error, variance) between two states—actual and desired—that we judge to be useless or harmful and wish to eliminate or reduce as much as possible. Resolving the problem means eliminating the gap, which in fact implies designing the control system so that it identifies the lever [X] that, by producing the results [Y], taking into account the states of nature [D], eliminates the gap with respect to the desired value [Y*].

Naturally we must first verify if the problem (gap) is real—that it derives from an actual, undesired state—or imaginary—that it derives from an unrealistic objective. If we specify the proper elements in the model in Fig. 5.18 we can obtain some interesting types of problems.

(a) *Problems of attainment* and *restoration* (or *maintenance*). This classification concerns the type of objective; if the objective is $Y^* =$ "a future state we want to reach" through alternatives that guide the present state toward that objective, then the problem is to produce the *change* that eliminates the gap, thereby *attaining* the objective. What stock should I invest in? Should I introduce a new product? Should I replace the machinery? Which TV channel should I watch tonight? If instead the objective is an historical state we wish to *restore* by bringing the present state back to the previous situation, or if it is the present state that we wish to maintain despite external disturbances, then the problem is one of *restoration or maintenance*. I have to change my tire, but how do I do this in the dark? The sink is stopped up and I desperately have to wash the dishes. Product quality is deteriorating and I have to block deliveries. *Attainment* problems are typically problems of result; *restoration and maintenance* problems are typically problems of control.

(b) *Isolated problems* and *recurring problems*. The distinction here concerns how often the problem comes up. The former occur in a restricted context and/or come up only once and/or can be quickly and definitively resolved. Recurring problems arise several times, are linked to previous problems, and their resolution each time influences subsequent problems. How often should I go to the hypermarket? I'm stressed out; I need something to calm me down. These are typical recurring problems because they can arise every day, week after week. Many species are becoming extinct. What can we do? This can also be considered a recurring problem since, for example, if we try to avoid the extinction of a species (panda) the problem arises for other species (whales, monkeys, parrots, etc.). It is not always easy to distinguish between the two types of problems; of course no problem is completely one-off, just as no problem is definitely recurring. We have to assess whether or not to maintain or break the links with other problems.

(c) *Problems which are deterministic*, *probabilistic*, *and uncertain in nature*. Here the distinction refers to the type of knowledge of the states of nature that can emerge. The problem is in a deterministic context if only one state of nature is foreseen for each alternative, so that for each X we can calculate the result obtained in terms of Y and choose the X necessary to produce $Y = Y^*$. Technical, engineering or simple financing problems are deterministic in nature. In many circumstances, for each alternative we can foresee alternative events whose probability of occurrence can be estimated; for each X we can calculate Y as the expected value of the random variable of possible results, thereby choosing X based on criteria that take into account the result and the *risk* associated with the attainment of the various possible values of Y. These problems are probabilistic in nature and are typical of the biological and social fields.

If alternative events are foreseen but the probabilities of their occurrence are not known, then the problem is *indeterminate* and the choice of *X* must be based on the results estimated with diverse criteria that take into account the *preference for risk* (for example, the *min–max* criterion or the *regret* criterion), caution (the *max–min* criterion), or absolute uncertainty. These problems frequently occur in all aspects of our daily life (what road to take to avoid traffic?) as well as in the physical, biological, and social worlds.

(d) *Problems with the states of nature* or *with rational adversaries*. When the values of the action variables *Y* do not depend solely on the *X* levers or on "natural" outside disturbances but instead are conditioned by countermoves by rational adversaries, then the problem can be compared to a game and be dealt with using decisional criteria from *game theory*.

From this distinction it is clear that problem solving is identified with a control system when it typically deals with *maintenance* problems, *recurring* and *recursive*, that are deterministic or probabilistic in nature. Even when problem solving deals with attainment and "one-shot" problems, which do not have repetitive decision-making cycles, it is still a control system in which the decision must lead to the elimination of the gap—and thus to a solution to the problem—in one cycle only. In these circumstances the problem-solving process is also known as the decision-making process (Harris 1998).

Problem solving is not only an application of the logic of control systems; it is a very complex process that utilizes a corpus of methods and techniques considered effective for solving problems in the various contexts we operate in as individuals, members of an organization, or as organizations.

> The work of managers, of scientists, of engineers, of lawyers—the work that steers the course of society and its economic and governmental organizations—is largely the work of making decisions and solving problems. It is the work of choosing issues that require attention, setting goals, finding or designing suitable courses of action, and evaluating and choosing among alternative actions. The first three of these activities—fixing agendas, setting goals, and designing actions—are usually called problem solving; the last, evaluating and choosing, is usually called decision making (Simon and Associates 1986, online).

5.9 Problem Solving and the Leverage Effect

After this brief survey I would like to reflect on the important contributions systems thinking and control thinking can make to understanding the logic of problem solving, beginning with a particularly important aspect.

If we represent problem solving as a control system, it is clear that, formally speaking, the problem appears as a "gap" between the objective and the present state we wish to modify through the choice of appropriate decision-making levers; in fact, *the problem is a symptom*, but of what?

We must search for the answer in the basic logic of systems thinking according to which every state, event, or situation must be viewed as the result of the action of

some system operating on one or more variables. By following this rule of systems thinking we are warned that a problem must not be identified with the *evident symptoms*—that is, the perceived gaps in the values of some variable—that require urgent measures: the *symptom* is not "the" problem because "the" problem is in the *structure* of the entire system and its whole dynamics. In order to solve "the" problem it is not enough to remove the *symptom*—settling for short-term symptomatic solutions (Sect. 5.2) with equally short-term effects—but to intervene on the *structure* producing that symptom.

> The definition of a problem and the action taken to solve it largely depend on the view which the individuals or groups that discovered the problem have of the system to which it refers. A problem may thus find itself defined as a badly interpreted output, or as a faulty output of a faulty output device, or as a faulty output due to a malfunction in an otherwise faultless system, or as a correct but undesired output from a faultless and thus undesirable system. All definitions but the last suggest corrective action; only the last definition suggests change, and so presents an unsolvable problem to anyone opposed to change (Brun 1970, online)

Employing the logic of control systems, we can redefine a problem as an anomaly, inconvenience, or difficulty in *carrying out systemic processes* or in *achieving given programmed objectives* which the manager of some management control system perceives when the control system displays an error, $E(Y) = Y^* - Y$. Systems thinking goes deeper and teaches us that a problem concerning the values of the variable Y arises precisely because of the bad—or undesired—functioning of the *generating system* of Y (normally considered a black box), which disturbs the optimal state, Y^*; the overall action of these disturbances can be considered the result of the action of external events. In order to solve the problem—that is, to eliminate the gap—we must reconstruct the *generating system* of Y (transform this into a white box) and determine the systemic levers, that is, the courses of action X, that operate to produce the desired Y^* in spite of D.

In other words, the consequences of each course of action do not derive from some *immediate* effect $X \rightarrow Y$ but from the overall *lever effect* ($X \rightarrow$ Generating System $\rightarrow Y$) that intervenes on the *causal map* of the system to control and on the *loops* on which the systemic lever operates. Thus the *definitive solution* of a problem almost never comes from variables directly connected to the symptom but from variables that are often distant, located along causal chains and loops in the generating system of Y; these variables must be recognized and their connections accurately described.

The model in Fig. 5.18 must be completed as shown in the model in Fig. 5.19.

The model demonstrates the important general principle that derives from the foundations of systems thinking we have just examined: we must not limit ourselves to considering problems only as undesired *symptoms* of *immediate causes* that are to be uncovered and eliminated (symptomatic solution), but rather as the undesired effects of the functioning of some *generating system* that must be recognized, specified, and controlled by means of control systems provided with appropriate systemic levers (definitive solution).
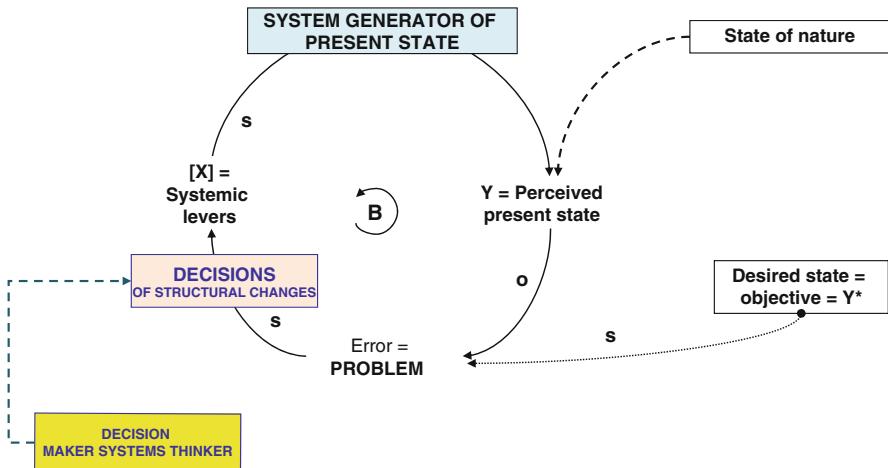
Fig. 5.19 Problem-solving model and the system generating the problem-causing state

Peter Senge, for that matter, has been clear on this point. He defines a *definitive solution* that exploits the potential of the system structure and its loops—not limiting itself to symptomatic interventions on individual variables—as the *leverage effect*.

> The advantage of systems thinking derives from the leverage effect—seeing in what way the actions and changes in the structures can lead to long-lasting, meaningful improvements. Often the leverage effect follows the principle of the economy of means, according to which the best results do not come from large-scale efforts but from well-concentrated small actions. Our non-system way of thinking causes significant specific damage because it continually leads us to concentrate on low leverage effect changes: we concentrate on symptoms of higher stress. We correct and improve the symptoms: but such efforts are limited, when things go well, to improving short-term factors, while worsening the situation in the long run» (Senge 1990, p. 131).

Obviously, as the quote clearly demonstrates, and Fig. 5.20 shows, exploiting the *leverage effect* means identifying the *loops* in the structure of the system—that is, the subsystems—which allow for *greater beneficial effects* on the symptoms, with a *minimum use of resources*, taking into account the time needed for the *leverage effect* to come into effect.

In fact, the causal diagrams do not automatically *highlight* any solution; instead, they must be carefully studied in order to identify one or more linked *loops*—called *system* or *structural levers*—which, by being acted upon, can cause the *leverage effect*. This represents the most difficult moment; if we cannot recognize the *system lever* that can bring about the *leverage effect*, then no decision can be said to be consciously taken and no problem definitively resolved.

Once again the fundamental instrument for reaching a definitive solution is the *structural map* of the situation; in problem solving the construction of the structural map must *start from the detected symptom*, from the difficulty encountered, in order
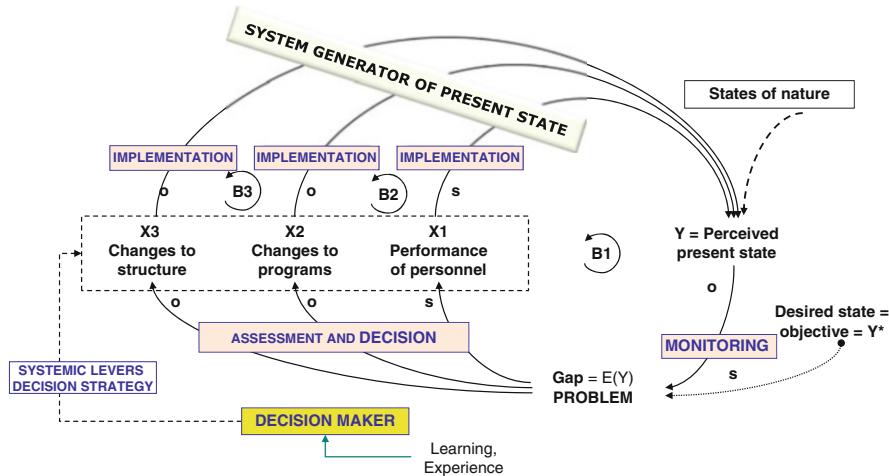
Fig. 5.20 Problem-solving model as a control system

to arrive, through successive connections, to the formation of the map of the processes. The problem emerges when we find anomalous relations or non-coherent trends in elementary processes or in chains of processes, which reveal unwanted reinforcements or undesired balancings, or when appropriate variable-generating programs are not applied.

The construction of the structural map makes it easier to *identify the alternatives* and assess the consequences, as shown in Fig. 5.20. There are three types of systemic levers [*X*] (that is, the action alternatives):

[X1]. Measures to increase the *performance of the personnel* that act through the physical structure of the processes that determine the dynamics to be modified.

[X2]. Measures that lead to *improvements in the action programs* (rules, routines, etc.) that activate the processes and determine their dynamics.

[X3]. Measures to change the *system structure*.

In order to identify the alternatives, the decision-maker must behave like the manager of a control system, evaluating the error and formulating *strategy* by adopting the well-known technique of "*what would happen if . . .; otherwise . . .*" Forecasting the consequences of each alternative means being able to simulate the changes the processes have undergone through their hypothetical implementation.

The structural map allows us to highlight the results obtainable in both the short and medium-long term. This *simulation* reveals all the usefulness of systems thinking; in fact, we must remember that every action has *immediate* consequences for the processes it modifies as well as long-term consequences, since every modification of given processes spreads to all the others linked to them in the structural map.

In order to identify the effective control levers the loops must be examined. In the medium term a *strengthening loop* can amplify, even unexpectedly, modest

effects that appear in the short term. On the other hand, *balancing loops* can render useless positive short-term results, even eliminating them in the long term. Finally, the chain of processes in the opposite direction can even create loops capable of inverting the direction of the expected results, so that the alternative we wish to introduce to eliminate a problem could actually worsen it.

Here are four important lessons from system thinking applied to problem solving:

1. If you want to solve a problem through alternatives that change the system's organization or structure, then redesign the map of the processes and analyze the loops; the new network of processes will permit an assessment of the risk that the consequences of the alternatives turn out to be the opposite of what was desired.
2. For a correct identification of the problem we must distinguish between process problems (often linked to operational programs) and structure problems; it is then necessary to evaluate the alternatives for structural changes through a loops analysis carried out over the medium-long term.
3. If you wish to solve a problem using alternatives that only modify the processes (thus employing new operational programs), pay attention to the alternatives that must produce an expanding or contracting dynamics, keeping in mind that, according to the Law of Dynamic Instability (Sect. 1.5), no expansion or contraction can be maintained for long; every reinforcing loop always produces a balancing loop which, in the long run, can change the direction of the dynamics.
4. In addition, be careful about the alternatives that eliminate an expanding or contracting dynamics; every balancing generates contrasting processes that can prevail and trigger the dynamics once again, often in the opposite direction.

This theoretical perspective which considers the solution of a problem as the search for the most appropriate control system to activate a process of change to modify the present state and bring about a future state (set as the objective) appears more useful the more frequently problems arise in *large systems* in which the *fundamental laws* of systems thinking are particularly evident and always act. Such systems are usually studied within the context of information systems and networks; however, their significance is more general, and they can be applied to various fields: social systems, economic systems, environmental systems, computer networks, logistics and transportation systems, large manufacturing factories, and so on (Dewan 1969), even if the concept has not yet been precisely defined.

> Regulation and control in the very large system is of peculiar interest to the worker in any of the biological sciences, for most of the systems he deals with are complex and composed of almost uncountably many parts.
>
> The ecologist may want to regulate the incidence of an infection in a biological system of great size and complexity, with climate, soil, host's reactions, predators, competitors, and many other factors playing a part.
>
> The economist may want to regulate against a tendency to slump in a system in which prices, availability of labour, consumer demands, costs of raw materials, are only a few of the factors that play some part.

The sociologist faces a similar situation. And the psychotherapist attempts to regulate the working of a sick brain that is of the same order of size as his own, and of fearful complexity.

These regulations are obviously very different from those considered in the simple mechanisms of the previous chapter. At first sight they look so different that one may well wonder whether what has been said so far is not essentially inapplicable. This, however, is not so. To repeat what was said . . . many of the propositions established earlier are stated in a form that leaves the size of the system irrelevant. [. . .]

Before we proceed we should notice that when the system is very large the distinction between D, the source of the disturbances, and T, the system that yields the outcome, may be somewhat vague, in the sense that the boundary can often be drawn in a variety of ways that are equally satisfactory. (Ashby 1957, pp. 244–245).

5.10   Complementary Material

5.10.1   Multi-criteria Decision Making

In order to offer an example of the decision-making process in the simple case of a multi-criteria decision under conditions of certainty (see Bouyssou et al. 2006), let us consider choosing a car by trying to satisfy the following five decision-making criteria, which we can also view as the "objectives" to achieve.

o_1—minimum purchase cost
o_2—minimum operating cost
o_3—maximum speed
o_4—maximum comfort
o_5—maximum representativeness as status symbol

Taking account of these decision-making criteria, we shall limit the choice to five models, which represent the courses of action.

c_1—two-seat sports car
c_2—four-seat sports car
c_3—big-engined station wagon
c_4—diesel station wagon
c_5—minivan

We must now specify the degree of preference for each alternative for each decision-making criterion (or objective).

– For the first three objectives we use the car manufacturers' sticker prices.
– For the objectives of comfort and representativeness we devise a score on a scale of 0–10, taking into account as well the "opinions" of the family members.

The *decision-making problem* can thus be represented by the *preference matrix* in Table 5.1. Written in the matrix squares for each line is the following:

Table 5.1 Preference matrix

	o_1 = purchase cost €/000	o_2 = operating cost €/km	o_3 = speed	o_4 = comfort	o_5 = representati-veness
c_1 – two-seat sports car	50	0.7	220	4	10
preference	2	3	1	5	1
c_2 – four-seat sports car	60	0.8	210	6	9
preference	4	4	2	5	2
c_3 – big-engined s-w	65	0.9	200	8	8
preference	5	5	3	2	3
c_4 – diesel station wagon	55	0.6	170	7	6
preference	3	1	5	3	5
c_5 - minivan	45	0.65	180	10	8
preference	1	2	4	1	4

– At the top center: the result for each alternative for each decision-making criterion
– In italics at the bottom left: the order of preference for the alternatives, on a simple scale of 1–5.

Observe that none of the alternatives (cars) is immediately preferable to the others; that is, there is no "dominant" alternative. The decision can be made using only a single decision-making criterion (objective) or all the criteria simultaneously. In the first case, the governance must specify the decision-making criterion adopted; in the second case, the *degree of relative importance* of the criteria.

This example immediately shows that every *multi-criteria* decision has no immediately determinable solution, depending instead on the *decision-maker's preference regarding the decisional criteria*. Every decision is made up of other decisions.

If we assume that the decision maker is "young" and not particularly sportive and assign to the decision-making criteria the weight indicated in the upper row of the preference matrix in Table 5.2, we can calculate a weighted preference index to determine, for each alternative, the *weighted average of the preference indicators*. These *preference indices* are shown in the last column on the right of the preference matrix in Table 5.2.

Since the scores from 1 to 5 are inversely related to the preference (index 1, bottom, preferred to index 5, top), in order to choose it is necessary to identify the course of action corresponding to the lowest average preference index (right-hand column); in our case, the choice would be to purchase the c_1 = "two-seat

Table 5.2 Preference matrix for a young decision maker

	$o_1 =$ purchase cost €/000	$o_2 =$ operating cost €/km	$o_3 =$ speed	$o_4 =$ comfort	$o_5 =$ representativeness	Index values
weight of objectives	*0.3*	*0.1*	*0.2*	*0.1*	*0.3*	*1.0*
c_1 – two-seat sports car						
preference	2	3	1	5	1	1.9
c_2 – four-seat sports car						
preference	4	4	2	5	2	**3.1**
c_3 – big-engined s-w						
preference	5	5	3	2	3	3.7
c_4 – diesel station wagon						
preference	3	1	5	3	5	3.8
c_5 - minivan						
preference	1	2	4	1	4	**2.6**

Table 5.3 Preference matrix for a "head of family"

	$o_1 =$ purchase cost €/000	$o_2 =$ operating cost €/km	$o_3 =$ speed	$o_4 =$ comfort	$o_5 =$ representativeness	Index values
weight of objectives	*0.4*	*0.2*	*0*	*0.3*	*0.1*	*1.0*
c_1 – two-seat sports car						
preference	2	3	1	5	1	**3.6**
c_2 – four-seat sports car						
preference	4	4	2	5	2	**4.9**
c_3 – big-engined s-w						
preference	5	5	3	2	3	**4.9**
c_4 – diesel station wagon						
preference	3	1	5	3	5	3
c_5 - minivan						
preference	1	2	4	1	4	**1.9**

sports car," which has an average preference index $= 1.9$. If we instead assume that the choice must be made by the "head of family," who assigns to the preference criteria the weights indicated in the matrix in Table 5.3, the choice would be to buy the minivan.

5.11 Summary

We have learned:

1. How to recognize or design the logical structure of a control system by recognizing the variables to control [*Y*], the objectives [*Y**], and the control levers [*X*] (Sect. 5.1).
2. There exist two control variants: structural and symptomatic controls, which intervene in many circumstances when we cannot identify the lever that acts directly on *Y* (Sect. 5.2).
3. How to evaluate the effectiveness and efficiency of a control system (Sect. 5.3) and how to strengthen *undersized* control systems (Sect. 5.4).
4. What the risks are of failure of the control carried out by a control system (Sect. 5.5).
5. What the risks are of failure due to the erroneous choice of levers. In this regard, two important archetypes were presented (Sect. 5.6): the FIXES THAT FAIL (Fig. 5.3) and the SHIFTING THE BURDEN (Fig. 5.4); a variant is also presented (Fig. 5.8).
6. We have analyzed four control pathologies: discouragement, insatiability, persistence, and underestimation (Sect. 5.7), and the four archetypes that represent them (Figs. 5.9–5.17).
7. Lastly, we have examined the *problem-solving* and *decision-making* processes and their relation to the control system (Sects. 5.8 and 5.9).

Part II
The Magic of the Ring

Chapter 6
The Magic Ring in Action: Individuals

In the arithmetic of life, One is always Many. Many often make one, and one, when looked at more closely, can be seen to be composed of many. (Lynn Margulis, cited by Habib 2011)

Let me briefly explain. Homeostasis is the tendency of a complex system to run towards an equilibrial state. This happens because the many parts of the complex system absorb each other's capacity to disrupt the whole. Now the ultimately stable state to which a viable system may run (that state where its entropy is unity) is finally rigid—and we call that death. If the system is to remain viable, if it is not to die, then we need the extra concept of an equilibrium that is not fixed, but on the move. What causes the incipiently stable point to move is the total system's response to environmental change; and this kind of adjustment we call adaptation. (Stafford Beer 1973, p. 54)

Part 1 presented the *theory* of control systems. A control system can be thought of as a *Ring* that rotates several times to guide the variables [Y], toward the assigned objectives [Y^*], through the control levers [X]. The first five chapters produced, so to speak, the anatomy of *Rings*, presenting their logical structure, their *modus operandi*, the main types, and the risks of failure in the control. In Part 2 we now move on to the *discipline* of Control, that is, training ourselves to observe, recognize, model, and simulate the action of the *Rings* in every micro and macro "environment" where they carry out their regulating function. It is not a question of overestimating the action of the *Rings* in our world but of realizing that the world itself is made up of *Rings* and that only their joint action can permit the birth and existence of living organisms, maintain the dynamics of the environmental variables, and produce evolution in nature as well as in society, collectivities, and organizations. It is no exaggeration to say that the *Rings* will not cease to amaze us and that we must recognize there is something *magical* in their existence. The *Ring* thus reveals its magic and becomes a *Magic Ring*. To accustom the reader to observing Rings, this chapter will take him on a mental journey through various "environments" characterized by the action of *Rings*. I do not claim that this treatment will be exhaustive; my objective is to indicate to the reader a path for recognizing *Rings* and their action which, through their *ubiquitous presence*, make all aspects of our world possible, in all its facets. By applying the first rule of systems thinking, which obliges us to *see the trees and the forests*, by zooming out

and in, we shall try to recognize control systems in ever wider contexts, beginning with the domestic and civic environment and then moving on to the physical and social environment in our world. This journey, though presented in summary fashion, reveals, in whatever environment we explore, an impressive quantity of *Rings* of various types and sizes. For this reason I shall put off examining the delicate control systems regarding societies and organizations until Chaps. 7 and 8.

6.1   Magic *Rings* Operating on a Wonderful Day

The simplest way to begin our journey into the "magic" world of control systems is to rethink our typical day. Obviously I will present some moments from "my typical day," but I invite the reader to reflect on his own experience, which could even be quite different. In any event, in both cases we will find the same *Rings*, though perhaps at different moments.

I am sleeping. My brain, in order to allow me to fall asleep and remain sleeping, has for hours raised the threshold values of perception, Y^*, so that all the modest sensory stimuli do not keep me awake. While sleeping, my *attention* is "turned off," so to speak. Suddenly the ringing of the alarm clock, Y, exceeds the threshold values, producing a variance between the nocturnal quiet and the noise perceived by my sense of hearing, which I perceive as annoying and which moves me to "reactivate" my attentive capacity (signs of awakening $= Y^* - Y$) in order to activate some control lever to make the annoyance cease. I move my arm, X, to reduce the distance, Y, between my hand and the source of the sound, Y^*, and I try to reach the alarm clock (distance from the alarm clock $= Y^* - Y$). Guided by the *routines* developed over many years, the tactile sensors of my fingers, X, reach the button, Y^*, after first having touched several objects on the night stand; once I recognize it (difference between button and the other perceived objects $= Y^* - Y = 0$), I press it (alarm position $=$ off) so that the ringing stops (alarm clock signal $= 0$). Thanks to these three *Rings* (reaching the source of the sound, recognizing the button, pressing the button) quiet is restored and drowsiness ceases. I am aware that describing the action of the various control systems by also indicating parenthetically the elements Y, Y^*, X, and E, which characterize these systems, weighs down considerably the text without adding further clarity; henceforth, when necessary I shall only indicate the type of control system (CS) that is at work.

Like a dove looking down from on high, as soon as I am awake I recognize the position of my slippers, and I quickly move my feet toward them (recognizing CS and attainment CS with a fixed objective). With my slippers on I assume a vertical position (dynamic objective), which is maintained by the control system for equilibrium (vestibular apparatus and leg and back muscles), and maneuver among the furniture in the room to reach the bathroom (tracking CS), where I begin a series of ritual activities that invariably are repeated each morning. After having completed the cycle of the Continuous Input-Point Output CSs (I shall leave this to the reader's imagination) I start the shower-shaving process. I shall not spend time describing the control system that allows me to shower but only mention that this is a

multi-objective *Ring* (control of temperature, water flow, amount of shampoo and soap). The point is to emphasize how complex the control process for shaving is and the fundamental role played in this by the CSs for the positioning of the blade on the face and the control of the razer's movement (attainment CSs). The fingertips, running over every millimeter of my facial skin, act as sensors of the control system to recognize the imperfect shave and the presence of residual whiskers to be eliminated. The action of these *Rings* can be more easily perceived by considering how many times the hand is taken away from the face and then once again reapplied, with extreme precision, and how many times the hand runs over the face again and again during shaving.

At this point the clothing and dressing ritual begins. The recognizing CSs enable me to choose the clothes to wear and to properly match my shirt, tie, socks, and shoes to my suit. The postural control systems allow me to quickly dress and to control, almost without reflecting on it (thanks to the muscle routines), the buttoning of my shirt and the size and position of my tie knot (qualitative CSs). Perhaps we do not reflect enough on how many *Rings* are activated to tie our shoes; we have to use both hands and a number of fingers as levers to move the shoelaces so as to form a knot with the proper tension.

There is no need for further detail about the problem of dressing; let us move now to the breakfast process, which is regulated by a number of repetitive control systems. There are many attainment CSs that allow us to grab our mug and spoon, pour the coffee, reach the sugar bowl, and pour the sugar, thereby eliminating the distance between the sugar bowl and mug while at the same time maintaining the spoon horizontal. The control system for the coffee's temperature leads us to activate the levers needed to bring the temperature to a level where we can drink the coffee without burning our lips. Several qualitative CSs are set off that enable us to regulate the quantity of sugar and cool down the coffee, without making the coffee too sweet or too cold.

Once breakfast is over the complex activity of making my way to work begins. After activating the recognizing CS to verify the elevator has arrived (in the contrary case, pushing the button to call it) I go down to the garage. In order to get in my car and leave the garage I activate a number of *Rings*. An attainment CS quickly operates several times to allow me to leave the garage by detecting the distances between the car and the right and left walls of the garage door (alternating objectives CSs). Then, after activating the recognizing systems for the presence of obstacles and pedestrians, I drive to my office. As we know, driving requires the activation of multi-objective and multi-lever CSs that entail the contemporary activation of a number of elementary *Rings*. While controlling the speed and distance of the cars ahead (variable objectives CSs), I recognize the signals from the traffic light and the cars ahead, accelerate, brake, turn the steering wheel, observe the lights, accelerate, brake, turn the steering wheel, brake suddenly, turn the wheel, stop, start moving again, turn left, gain speed, etc. This, in short, is how all the *Rings* needed to drive to my final destination act, dozens of times per minute. And let us not forget that usually the control systems of stress and anger are also activated; these are also needed to arrive at the office and start work under the

proper psychological conditions. There are dozens of people to recognize, greet, accompany, and dozens of documents to recognize, read, file away, or respond to that require the activation of a multitude of recognizing CSs as well as the necessary attainment CSs. Likewise, it is easy to see how many attainment and recognizing CSs are needed to write at the computer.

All the activities so far described obviously require the operation of systems of biological control that make up our body, in particular all the CSs of vision, hearing, tactile capacity, as well as postural systems that modify the position of our limbs to allow us to remain balanced, walk, drive, etc. I shall let the reader imagine, by referring to his personal experience, the number and type of *Rings*, of the most varied nature, that are activated at other times during the day, when we speak with our spouse, children, friends, read the paper or relax in front of the TV.

Now it is nighttime. In bed, automatic mental control systems slowly deactivate our attention while also raising the threshold levels to inhibit the perception of sensory stimuli. Our eyes close and we start to fall asleep (non-REM followed by the REM phase), in anticipation of starting a new day, one of many thousands of future days in which millions of control systems will act, second after second, billions of times, to make our existence possible.

Now that we are aware that our lives as individuals would not be possible without the action of these many *Rings*, we can certainly realize, perhaps with some amazement, why we must speak of "magic" *Rings* as the basic instruments of our existence.

6.2   *Rings* Operating in the Domestic Environment

The previous *section* took the simplest route to begin a journey in the universe of the most evident and verifiable control systems which, day after day, allow us to continue down our path in life. However, the *Rings* are everywhere and make possible our behavior in life. The *Rings* do not only allow us to exist as individuals but operate in the most unthinkable places, situations, and circumstances; the fact we are not used to noticing them takes nothing away from their ubiquity. To accustom ourselves to the Discipline of Control it is useful to proceed by means of progressively zooming in or out to observe our environment: start with the simplest and most evident control systems—which occur daily in our home, work or town environments—and then extend our observation and reflection to those that operate at a wider or a narrower level.

The domestic environment is ideal for our first zooming activities regarding this discipline. Our household appliances are by now all built using automatic control systems. The fridge maintains the temperature objective when there are changes in the internal conditions (introduction of hot food) and external conditions (the room temperature increases). The microwave shuts off when the programmed time objective is reached, or when we open the door.

Along with automatic systems, where our intervention is limited to setting the objective—as in the case of the elevator in Fig. 2.19, where we simply push the button to indicate the objective that the "machine" then achieves without any further intervention on our part, or the refrigerator or air conditioner, where the objective is set through the appropriate analog or digital instruments—many others exist where we ourselves are components of the control process, either because we monitor and observe the objectives (temperature, levels, etc.) or because we regulate them (rotation of the levers, pressure on buttons, opening of doors and windows, etc.).

We become part of the control process at sundown, for example, when, no longer able to read our newspaper—despite the fact the control system for the eye has dilated the pupils as much as possible—we regulate the light in our room to maintain a visibility limit, turning on artificial light or opening wider the shutters. However, today on the market we can buy sensor devices to automatically regulate the light throughout the day, after setting the objective for the desired amount of light. Every time we pour wine into a glass (halt control system) or control the stock of food in the fridge (PI–CO control system), we are at the same time monitors (level of wine and stocks) and regulators (stop pouring and visit the hypermarket).

Our computer is a galaxy of control systems in terms of both the hardware and the various software we use; however, it could not function without our intervention; the same occurs with our gas stove, electric oven, air conditioner, television, washing machine, dishwasher, iron, radio alarm clock, and wrist watch. And what about the flow controls we daily carry out using the regulators for our various taps to control water levels, flows and temperature? How can we cook a cake or make risotto if we do not regulate the level of the gas and how long it is on? Today ever more versatile kitchen robots are becoming common and android robots are already on the market, which represent true ambulatory control systems. We can also consider the control of our pantries and how we change our way of dressing with the changing seasons, the places we frequent, and the circumstances in which we present ourselves. In short, our entire domestic life is only possible thanks to the many *Rings* that permit us to achieve objectives of well-being or to respect the myriad of work and safety constraints.

In all domestic control systems, we play a principal role as users and managers–governors, and only our experience enables us to successfully achieve our objectives. Grandma's cake never burns because she checks how it is cooking minute by minute, while it is likely her granddaughter burns her cake from being distracted by the endless number of telephone calls she receives.

However, I can safely state—referring to Fig. 6.1, which derives from Fig. 3.1—that our domestic life is characterized by continual control processes in which, in varying ways, we play a role, either in a minimum way [①= = we only set the objectives], an intermediate manner [①+②+③= = we set the objectives, note down the error and decide the action to take on the control levers], or in global terms [①+②+③+④ = = we form the entire chain of control; we are the control systems]. If you are not convinced of this, "look around."
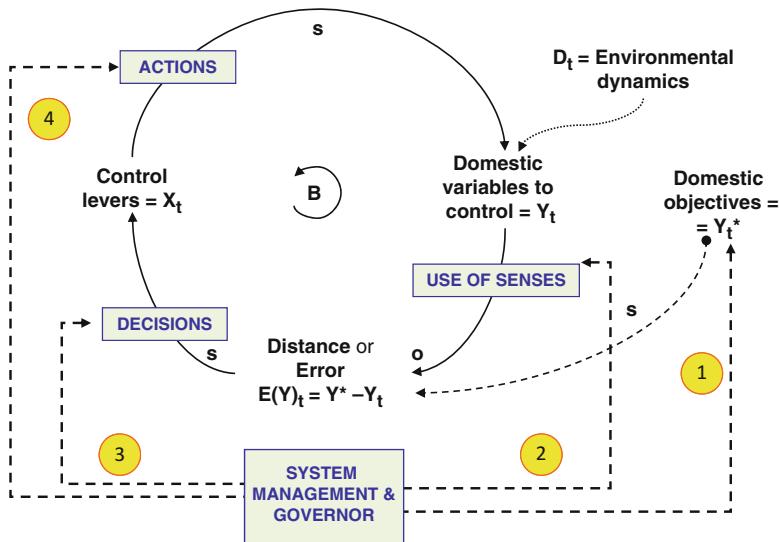
Fig. 6.1   General model of our intervention in domestic control systems

If we even consider our car as part of our environment, we can see that while driving it we form together a formidable multi-objective and multi-lever control system—of which we are the governors–users (we take a trip and establish the policy indicating the priority of the objectives), the manager (we decide the strategy and regulate the control levers), even the apparatus in the chain of control (we record errors, turn the steering wheel, and activate all the technical apparatuses)—that permits us to reach our desired destination by controlling direction, speed and travelling time, while taking into account speed constraints and external disturbances (e.g.,, pedestrians crossing the street).

If you are amazed by the actions of a "Formula 1" race-car driver, who controls his car (accelerator, brakes, and gears) to gain the maximum speed, without swerving, while paying attention to the other drivers, do not underestimate your own abilities when, each morning, you go to work by car in the crowded city streets. The control systems you manage are more complicated than those of the professional driver, having to take into account traffic lights, pedestrians, bicycles, cars braking suddenly, etc., which force you to regulate your speed, lateral and frontal distances, curves, lights, look out for the traffic police, and so on, through continuous actions on the available levers. The car has hundreds of automatic control systems, most of which we ignore (me included), that allow us to drive without worrying about minute objectives and constraints that must be met by the various processors controlled by the automatic electronic devices in the propulsion mechanism, the mechanical structure and the electrical system. In many car models today, control is becoming almost total: the lights turn on automatically when luminosity is reduced; the wipers start automatically when it rains and automatically go faster as the rain intensifies; all kinds of sensors tell us when the car is in
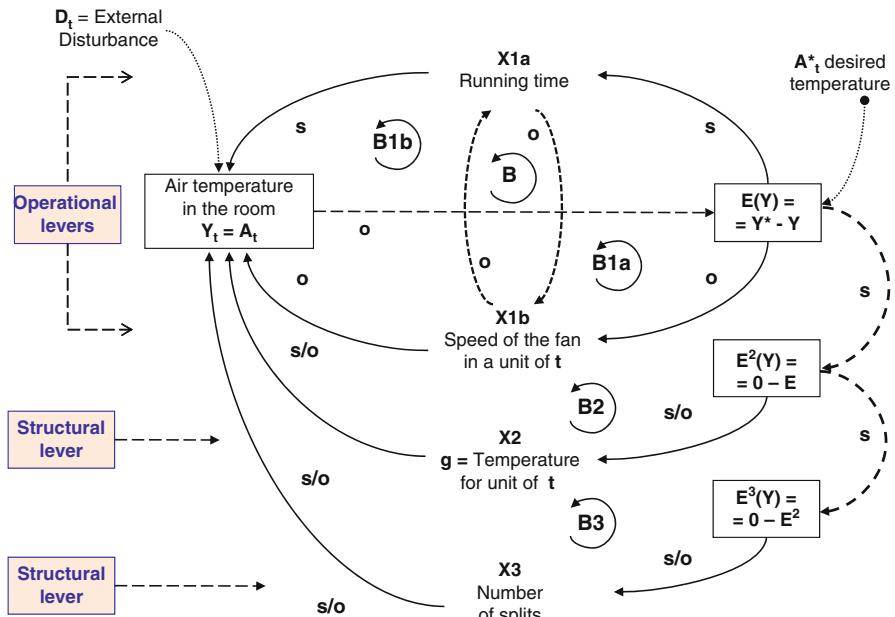
Fig. 6.2 Air conditioner as a three-layer control system

gear, the doors are open, the tire pressure is low, what the petrol level is or the maintenance requirements are, and so on. Some models are equipped with systems to control (by preventing) nodding off by the driver, by detecting the time the eyelids are lowered and producing vibration in the steering wheel or some other algedonic signal. Recent models reduce speed—without the driver's intervention— when there is an obstacle in the road and even park automatically.

For the most part domestic control systems are multi-lever and multi-layer. Even the air conditioner achieves its temperature objective using many levers. The model shown in Fig. 6.2, which completes Fig. 2.6, involves a three-layer system.

In the first layer the temperature is controlled by two operational levers: the amount of time the air conditioner is on and the fan speed (if available). If these two levers are not enough, there is a structural lever available that modifies the action rate to increase the temperature of the air that is expelled. And if this too is insufficient, recourse can be made to a third layer, which increases the number of splits per room.

What is more interesting, however, is observing how many control systems, even functioning as autonomous systems, are interconnected—or interconnectable—so as to form control systems that are much more powerful and flexible in achieving their objectives. The control of the hot (radiators) and cold air (air conditioners) interact to maintain a comfortable temperature in our homes year round. Yet even the control of the heat emissions from our appliances or our lighting system, or the control represented by the opening of our windows in summer and winter contribute to maintaining a comfortable temperature.

We have thus arrived at an initial conclusion: though not always possible, when we set an objective in our domestic life we can almost always "put in play" a multitude of *Rings* that form an integrated holonic super system of control, whose various component *Rings* represent the operating levers (Sect. 3.8).

It is not farfetched to state that the first forms of human progress at the dawn of civilization were precisely the result of the continual search for control processes and systems that make life possible; avoiding burns, producing and preserving food, finding protection from dangers, and so on, provided the stimulus for controlling fire, raising animals, growing crops, and building homes. Since then man has come a long way, but even today the search for a comfortable life continues with many projects to aid man in his domestic chores. We can foresee that the continued progress in *domotics* will before long allow for the *remote control* of all our household electrical appliances, which will be equipped with automatic control systems to carry out their processes and limit human intervention, reaching the point where objectives are also set in the form of *standard procedures*.

Recognizing CSs are also part of electrical appliances. For example, vacuum cleaners become automatic robots that explore every angle in the house; washing machines recognize the weight and type of load and choose and measure out the detergent to use as well as the washing time. And recently a brand of refrigerator has been marketed whose models are equipped with ten sensors for the automatic and dynamic control of the temperature and humidity levels to permit the best preservation of our food by signaling the food that will expire next and that already past the expiry date. The "virtual chef" is a multi-objective recognizing, attainment and halt control system that simultaneously regulates the many cooking variables for a variety of food. There are even *safety doors* available, equipped with sensors that display the temperature, which are automatically blocked when the error indicates a dangerous temperature.

Progress in computers has led to the creation of control systems to aid people with physical problems. "Emc" and "Paravan" now sell a "four-way" joystick that encapsulates the three basic control levers needed to drive a car. Through a single lever the joystick manages the steering wheel, gas pedal, and brakes, and it can be used by disabled people for driving. Soon it will probably also be equipped for all other drivers, since we all would like to control all the levers with the minimum fatigue and maximum safety when behind the wheel. Already commercially available are apparatuses that use voice commands to regulate lights, directional indicators, windshield wipers, horns, door locking, and convertible tops. From what we read in many newspapers and on Web sites, progress in control systems for car transport does not stop there; there are now even completely automatically driven cars that arrive at the programmed destination without the need for a driver.

The *Google driverless car* is a project by Google that involves developing technology for driverless cars. . . .The project team has equipped a test fleet of at least ten vehicles, consisting of six Toyota Priuses, an Audi TT, and three Lexus RX450h, each accompanied in the driver's seat by one of a dozen drivers with unblemished driving records and in the passenger seat by one of Google's engineers. . . The system drives at the speed limit it has stored on its maps and maintains its distance from other vehicles using its system of sensors. The system

provides an override that allows a human driver to take control of the car by stepping on the brake or turning the wheel, similar to cruise control systems already found in many cars today. …In August 2012, the team announced that they have completed over 300,000 autonomous-driving miles (500,000 km) accident-free, typically have about a dozen cars on the road at any given time, and are starting to test them with single drivers instead of in pairs. Three U.S. states have passed laws permitting driverless cars as of September 2012: Nevada, Florida and California (Google/Driverless, car project 2012, Online).

The engineers from Google explain that the Prius utilizes a series of cameras and a roof-mounted, spinning laser to see what is going on around it. The result is a vehicle which might just be safer than one with a human behind the wheel. However, according to the report, the goal of the system is not to completely remove the driver from the equation; the system is pitched as more of a "super cruise-control" than a full auto-drive system. The theory is that it would be useful for traffic-filled commutes to and from work, and it might be a nice solution to eliminate or reduce distracted-driving. Get a phone call? Hit the Google button and let the car have the wheel while you take your call. (Google/Prius 2010, Online)

Remaining on the topic of control systems for transport, for some years now a very sophisticated instrument known as "Segway" has been sold for individual, low-speed locomotion. "Segway" is a bicycle with two parallel wheels equipped with multiple MEMS, that is, solid gyroscopic sensors that can emulate human equilibrium and which permit both control of vertical stability for the passenger, even when the vehicle is stationary on two wheels, as well as total maneuverability, with forward and backward direction, since the "Segway" can easily navigate curves and attain speeds of around 20 km per hour. "Segway" can, in all regards, be considered a multi-lever and multi-objective control system, since stability is maintained through electric apparatuses that eliminate even the minimum error in equilibrium detected by the multiple sensors it is equipped with. Change in direction comes from shifting a vertical, inclinable steering column right or left, as if this were a man-sized joystick.

As a final example of automatic systems in the domestic environment, in Italy the Computer Vision Laboratory, Visilab, at the University of Messina, has created software for a blind assistant.

Blind Assistant is an attempt to realize a software that can help men and children that are blind or visual impaired. This software doesn't require a complex and dedicated hardware: in fact, it is designed to run on simple and cheap devices such as Sony Playstation Portable (PSP). Blind Assistant uses Nanodesktop technology for windows environment, it is able to speak with the user using the speech synthesis engine ndFLite, and it is able to recognize the commands spoken by the blind through the speech recognition engine ndPocketSphinx. In this way, the blind can control the program using the integrated voice interface. At the moment, Blind Assistant provides 7 functionalities: Face recognizer, Room recognizer, Optical chair recognizer, Color scanner, Mail reader, DMTX scanner, Clock/Calendar. The face recognition technology is provided using ndOpenCV, a porting of the original Intel OpenCV libraries that has been developed by the same author, and which is part of the Nanodesktop SDK. The system is able to recognize the name and the position of the people that are present in that moment in your room, using complex algorithms (such as Intel EigenFaces/PentLand PCA) [. . .]. (Visilab (2013, Online)

6.3 Overhead *Rings* in the External Microenvironment

We need only leave our houses and look around to understand that we could not survive long in our outside microenvironment—represented by our city and region—without a very large number and variety of *Rings* that show us how to move and adapt ourselves in this environment. I would also observe that we ourselves represent direct control systems that carry out a large number of control processes: for example, whenever we avoid running into other passersby; judge the best moment for crossing the street, despite the traffic; check if we need to stop for the traffic light; avoid puddles; reach our destination; and select the bell to ring to gain entrance. Moving around the city or countryside as control systems, we can achieve our quantitative and qualitative objectives and abide by the large number of constraints that limit our discretionary power.

However, the outside environment is rich in various kinds of control systems which we are not even able to perceive. Many of these are automatic, while others require the intervention of man (teams of operators), which are carried out above our heads, as silent as they are useful: these are *overhead* control systems.

Many essential *overhead* control systems are managed by *policy makers* and *public regulators* to provide public services or protect public health, in order to satisfy a multitude of stakeholders that make up the collectivity in a given territory. Figure 6.3 shows the general model of how overhead control systems operate in the external microenvironment to satisfy the needs of stakeholders through the provision of services meeting given quality standards. Among the many overhead control systems that exist, I shall only briefly mention the most useful ones.
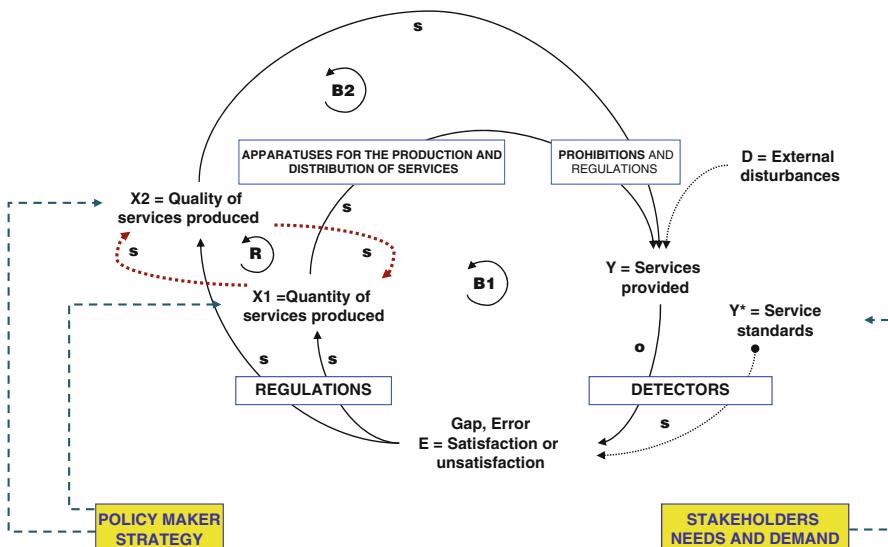
Fig. 6.3 Overhead control systems operating in the external microenvironment
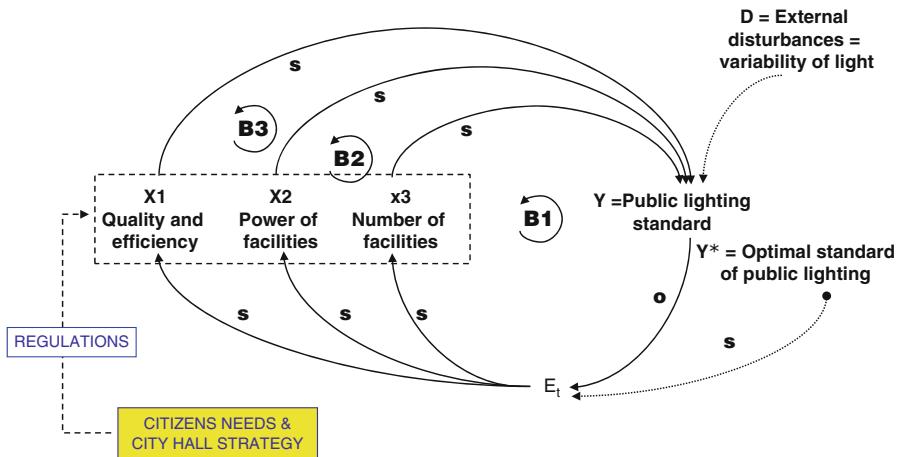
Fig. 6.4   Control of the level and quality of public lighting

We do not realize how many controls are needed to supply that extraordinary good, that is, the energy to illuminate streets and houses and to operate all the mechanisms that could not work without it. Whether or not it is produced through hydroelectric processes, some kind of combustion process, wind or solar power, or even nuclear energy, the production of electricity must be minutely controlled in all its phases in order to maintain constant the tension and ensure there are adequate flows by automatically importing energy from surplus producing countries to avoid blackouts during periods of intense use. Generally this involves automatic engineering controls on the electrical distribution network that, in any event, refer to well-known systems.

A multi-lever and multi-layer control system we are not aware of is the one that guarantees regularity in the level of public lighting, with the automatic switching on of street lights when the outside light falls below a specific threshold and the turning off of the lights in the opposite case, as shown in the simplified model in Fig. 6.4.

Parallel to the control system for the quality of public lighting is the control system for *energy savings*, which determines the optimal period for turning on the lighting facilities in streets, squares, public buildings, stadiums, etc. Lastly, let us not forget the control system that allows us to pay for our energy use through meters and, in the event the meters do not permit an automatic reading (detection devices), personnel trained to read them (detection). The bills are then sent out and the payment verified.

What we have just observed about the production and distribution of electricity is equally true for *natural gas* for domestic or industrial use. We do not realize how many controls there are, not only in the initial phase when the gas is injected into the distribution network but also, and above all, in keeping the network working perfectly efficiently. A qualitative control system placed in the switchboard system in the main junctions of the distribution network allows maintenance teams to
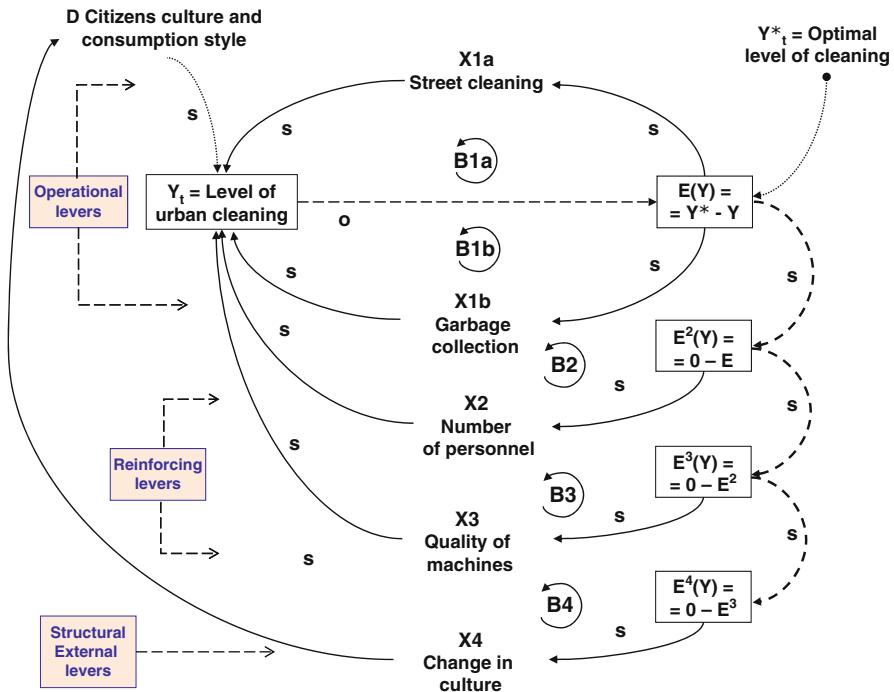
Fig. 6.5 Control system for solid urban waste

promptly intervene (control implementation) as soon as a malfunctioning is detected (detection devices) with respect to the set standards; our noses can even detect malfunctioning with respect to set standards when there is a gas leak in our town.

Residential centers could not exist without the highly important control systems for liquid and solid wastes, which even the Romans were aware of: before any urban expansion they enlarged the sewer system. Today the production and distribution system for the necessity goods in life is also a powerful producer of solid wastes of all kinds that must be eliminated every day.

Sewers and systems for collecting and disposing of waste are vital for avoiding health catastrophes, but these systems must be guided by very sophisticated control systems.

We all know what happens when there is a malfunctioning in the control system that regulates the collection of solid waste: suffocating garbage piles disrupt normal activity in a city and lead to intervention by our man-run and civic hygiene control system that forces us, despite our protests, to become active participants in the authorities' efforts to activate or strengthen the control levers for the volume of waste.

The model in Fig. 6.5 indicates the levers that can be used to control solid urban waste and for street cleaning.
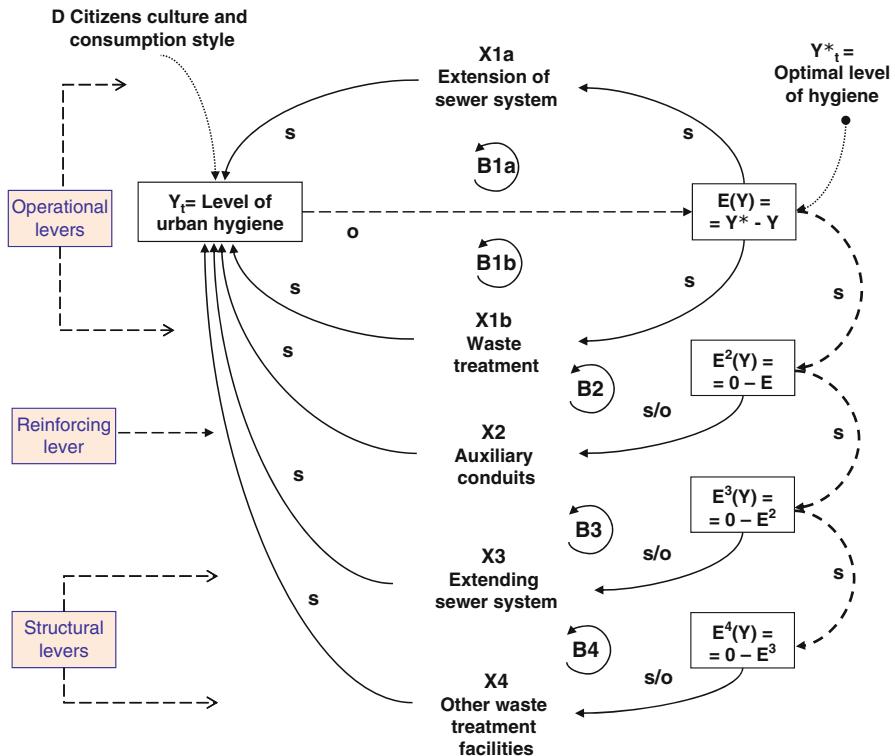
Fig. 6.6 System of control of liquid waste treatment

Figure 6.6 presents a model that illustrates how a control system for urban hygiene can operate through the sewerage system. In recent decades our cities have experienced another important problem: along with the elimination of liquid and solid waste, there is the problem of the elimination of *gaseous waste* from all types of harmful emissions. Particularly felt is the need to control *particulate matter*, or PM_{10}, which is so dangerous to our health because it is composed of a harmful *mix* of particulates, that is, solid and liquid particles suspended in the air (whose diameter is equal to or less than 10 μm) that contain dust, smoke, soot, and aerosol (a mix of micro drops of various liquid substances).

Particulate matter comes from various sources, among which hydrocarbon combustion by cars and heating plants, fires, volcanic emissions, tire wear, asphalt particles, micro-pollens, etc. It is not yet possible to control pollution levels with automatic control systems. However, some well-known levers can be adopted for the control of particulate matter such as traffic restrictions, time and temperature constraints on home furnaces, incentives to adopt solar energy for heating, and so on.

To carry out the necessary controls with concrete measures large cities have installed a network of detectors for pollution levels that allow authorities to detect deviations from the tolerance limits (constraints), which can lead to corrective
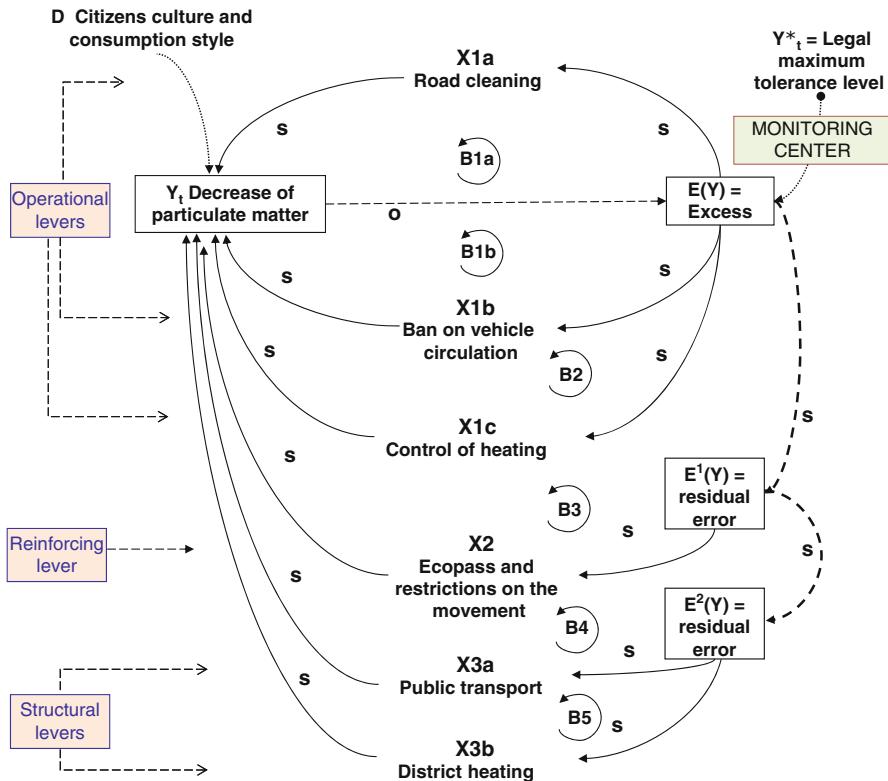
Fig. 6.7 Control system for fine particles

measures, as long as the control bodies, which are governed by bureaucratic and political structures, are willing to intervene. It is no surprise that orders from the regional authorities to limit traffic in a certain region are never accepted by all the cities in the region and that, for certain reasons, some local authorities permit a continual increase in the pollution thresholds.

Figure 6.7 illustrates a simple model of a control system for *particulate matter*. Note the multiple levers that are used and their classification into operational, auxiliary, and structural levers based on the environment in question and the regulatory organization.

I should mention two other types of pollution: pollution from the intense *magnetic fields* produced by electric energy distribution systems and by systems that produce communications signals, and *acoustic pollution*, caused by noisy instruments or traffic. We must unfortunately admit that, unlike other types of pollution, the control systems for these two forms are quite tenuous. Although there are monitoring instruments to detect levels above the danger thresholds, the control system has two weak links: the scarcity of control levers to act on the sources of pollution—structures solidly "in place" in the residential areas in

cities—and the lack of willingness on the part of public authorities to implement the few existing control levers, such as a ban on new plants, moving existing ones to other sites, or installing protective barriers.

A brief mention of a sore point: why can't we control the punctuality of *public* buses in large, crowded cities with choking traffic? The answer is twofold: on the one hand, there is a lack of effective control levers, and on the other the external disturbances to the scheduled route times are so great that control is quite problematic. What levers could the public regulators implement? More efficient route planning? The creation of preferential lanes? Making it easier for users of public transport to get on and off the vehicles? Further development of the public transport network? All of the above can be effective in theory, but in practice traffic disturbances thwart the efficacy of any lever.

The range of overhead control systems is not of course limited to the few examples above, but these nonetheless are sufficient to make the reader reflect on the urban environment and recognize there are many other *Rings* that make possible the world around and above us. In a few years' time we will have intelligent cities that integrate all the control systems we have so far examined—waste disposal, hygiene, transport, pollution, etc.—to form a global control system that operates thanks to the regulation of a powerful central information processing system. Here is an example of a successful experiment.

> PlanIT Valley will combine intelligent buildings with connected vehicles, while providing its citizens with a higher level of information about their built environment than has been possible previously. Its efficiency will extend into the optimum control of peak electricity demand, adapted traffic management for enhanced mobility, assisted parking and providing emergency services with the capacity to have priority when needed in the flow of traffic.
>
> PlanIT Valley will enable the enhanced monitoring of the vital signs of urban life, the condition and performance of vehicles and infrastructure. As a result, managers will be able to optimize normal daily operations of the city and provide greater certainty in reacting to extraordinary events through real-time modeling and simulation. With a view to incorporating new developments, urban management control systems will be updated with the latest information and technology as these emerge.
>
> Urban mobility will be based on lightweight and fuel efficient electric or hybrid vehicles built with the use of new composite materials. When connected to the city and services, this will make better use of mobile computing capacities in vehicles, benefitting from rich wireless capabilities optimizing the use of bandwidth. These will enable citizens to fully exploit energy storage capabilities while driving, use electricity more efficiently in homes and sell excess capacity to the electric power grid. (PlanIT Valley Project 2013, Online)

6.4 *Rings* Acting in the External Macro Environment

If we zoom out and broaden our observational horizon to the macro environment, we observe a large number of control systems operating in various contexts. However, it is not easy to identify these since we are not accustomed to thinking of macro phenomena as elements in some control process. With a bit of practice the
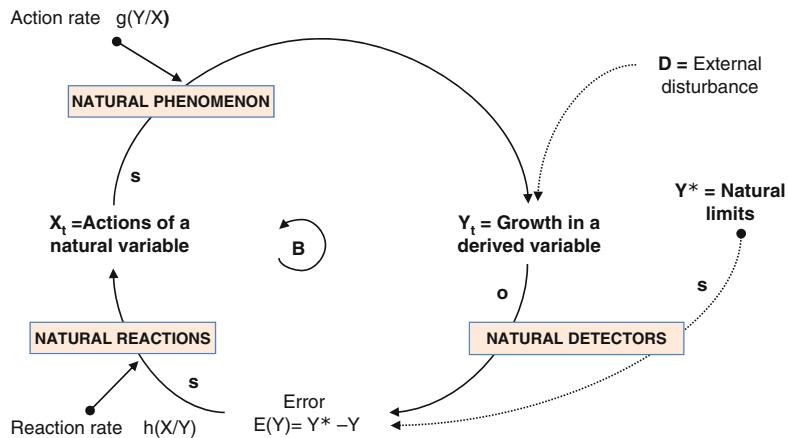
Fig. 6.8 General model of how "natural" control systems operate in the external macro environment

discipline of control will enable us to better understand the ubiquitous presence of *Rings* in the macro environment as well.

Let us above all consider the significance of many of the structures we see around us: dams, canals, boundary walls, and observation towers, all testify to man's desire to control his hostile surroundings. These are levers to dike a river or sea, or to protect ourselves from the enemy hordes. The more frequent the floods were, the more embankments that were built; the larger the invasions, the more widespread the ramparts and walls for protection, as shown by the Great Wall of China.

Beyond these evident levers for man's control over nature, we must recognize that on a macro level "nature" itself is regulated by the action of a large number of "natural" *Rings*. These systems are not always easy to identify and represent, but their action is constant; the discipline of control must accustom us to recognizing and constructing models of them. The difficulty in recognizing such *Rings* comes from the fact that environmental dynamics—unlike the case with man-made control systems—do not achieve explicit *objectives* but rather avoid exceeding certain natural "*constraint values*," or even achieve "*limit values*" believed optimal in relation to certain environmental conditions. Natural control systems always behave in the same way, but, independently of the peculiar nature of "natural" limits or constraints, they can be described using the usual control system model, as shown in Fig. 6.8.

When a system is in an equilibrium state, $Y_t = Y^*$, with respect to the value of a "natural" variable considered optimal, or binding, in those conditions, and this equilibrium is changed by the action of environmental disturbances—often considered random because of our ignorance of the causal factors that give rise to them—there then emerge variances, $E(Y)$, which, though not immediately perceivable, trigger changes in some other "natural" variable, X_t, which brings the system back to equilibrium, thereby eliminating $E(Y)$.

There are three main difficulties in perceiving a similar control action. The *first* is simply the fact we do not always know the optimal values, Y^*, that the "natural" *Ring* tries to achieve and maintain. The *second*, which is more "fleeting," is that the optimal values occur when particular environmental conditions exist in the form of specific values of other accompanying variables that are not always known or easily knowable; these are not part of the system but belong to the external environment. The *third* difficulty is connected to the fact that, because of the dynamic nature of the accompanying variables, the optimal values change over time, even rapidly, with the result that Y^* is constant only for a brief period and can vary, often unpredictably, in subsequent periods.

To demonstrate this let us take the example of a very simple "natural" *Ring* we all can observe: the control system leading to the formation of "fair weather small cumulus clouds," those isolated small clouds in the shape of a small spherical pile that, during the warm days of spring or summer, suddenly appear in the sunny sky above the plains, above the seas, or hovering over the mountain tops.

These small cumulus clouds can be interpreted as a *symptom*—that is, the error $E(Y)$—that occurs during a control process under way (invisible to us) that leads to a re-equilibrium in the density of small masses of air (commonly called "bubbles") through water vapor condensation, when the latter reaches a certain density at a given atmospheric pressure with respect to the surrounding air.

Referring to Fig. 6.9, we can briefly define the density, $D_t = H_t/T_t$, of a mass of air as a ratio between the humidity (H_t) and temperature (T_t) at a certain pressure. The higher the humidity and the lower the temperature, the more the increases in density.

We can imagine there exists a limit value, D^*, at which the mass of air is saturated and gives off the humidity it contains, at the same time lowering the temperature. D^* can be interpreted as a limit (or a constraint/objective) of the control system that produces small fair-weather cumulus clouds.

With the aid of Fig. 6.9—which is flipped over with respect to the general model in Fig. 2.2, in order to visualize the small cumulus cloud phenomenon—we can easily describe the process.

The irradiance increases the temperature (T) of an air mass that persists over an area with a particular shape, such as an area of heat-retaining *dark land* which is surrounded by light-colored land, resulting in the air mass heating up more than the surrounding air. At the same time, the irradiance favors evaporation that increases the humidity (H) of that same mass on the dark area. In most cases the increase in temperature is greater than that in humidity, so that the density $D_t = H_t/T_t$ decreases and the air mass becomes lighter, following the direction "**o1**."

Nothing happens until a chance event (external disturbance)—for example, a gust of wind, the movement of air from a passing freight train, or some other event—detaches the air bubble, which, being light due to the high temperature, due to Archimede's buoyancy force, tends to rise and increase its distance from the ground (direction "**o2**"). Thus we can think of *temperature* and *humidity* as two control levers (X1 and X2) in the process that involves the variation in air density $Y_t = D_t$). Rising toward cooler regions, the bubble cools, thereby increasing (direction "**s1**") the density (D_t). This is the first phase in the control process.
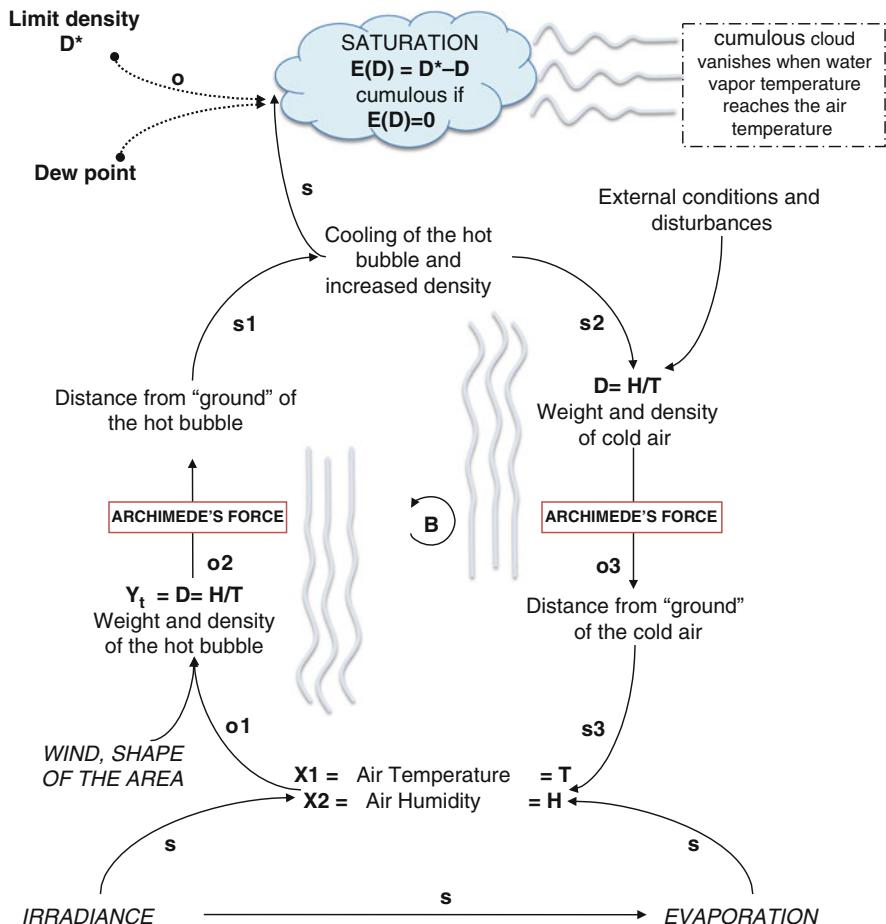
Fig. 6.9 The formation of fair-weather cumulus clouds and rising thermals

We assume that at a given temperature, called the *dew point*—which varies according to a number of conditions—the density, (D_t), reaches the limit density, (D^*); the air bubble then becomes saturated and releases its humidity, which, as it condenses, becomes visible as a "fair-weather cumulus cloud" that releases its "latent heat from evaporation," probably further reducing its temperature. However, the process is not over yet. Once the rising air cools and releases its humidity it regains weight (direction "**s2**") and sinks toward the ground (**o3**), where, if the conditions are those of the previous cycle, it heats up again and gains humidity (**s3**). If the control system repeats loop [**B**] for several cycles (there are 3 "**os**"), various air bubbles are produced that rapidly follow on one another to create a true column of rising hot air, a rising thermal, that comes up against a flow of descending cold air. It is these opposing air flows that are a source of enjoyment for lovers of the air sports of gliding, hang gliding and paragliding. Practitioners literally

jump from one cumulus cloud to another in the search for updrafts (avoiding downdrafts), which take them higher and allow them to remain flying for hours. I do not know whether Fig. 6.9 provides an effective explanation, but one thing is certain: cumulus clouds form and represent a visible reality.

This raises an *initial question*: what meaning do we attribute to cumulus clouds? As we can see from Fig. 6.9, they are not the output of some process but the visible *symptom* of the action of natural invisible *Rings* when the rising air bubble reaches the limit value, D^*. They are not the effect of the control levers but of the trends in $E(D)$. It is enough to rotate Fig. 6.9 180° to clearly recognize the structure of the general model of control systems we are all familiar with, in which the small cumulus clouds become the revelatory sign of the elimination of the error regarding the saturation of the air.

A *second question*: why don't the fair-weather cumulus brighten up all our spring and summer days, but instead are only rarely visible, more of an exception than the rule? The answer lies in the peculiarities of the control process, which depends on very particular conditions regarding irradiance, evaporation, wind, air conditions above the ground and, above all, on the dew point and the limit density value, (D^*), which vary according to external circumstances and can be considered as disturbance factors.

Moreover, the model in Fig. 6.9 can be considered a general one for interpreting the formation of clouds as a consequence of a convective heat transfer resulting from shifting air flows that are heated at ground level and pushed toward high-altitude atmospheric strata; the temperatures of these air flows falls by around 6.5 °C for every 1,000 m of altitude, as is well known to pilots flying at 10,000 m, where there is a risk of ice forming on the wings. All air masses heat up and grow moist during summer days, but if the hot air slowly and continually rises, if there is not much humidity, if the above-ground temperature is constant for hundreds of meters, and if D^* is very high, then the conditions do not exist for the air to release its humidity, and thus no cumulus will form. For this reason the "fair-weather cumulus" are a rare rather than a normal occurrence.

The explanation for the formation of cumulus clouds leads us to the common and general phenomenon of *cloud formation*. As the preceding process clearly indicates, clouds form in the same way: they are the result of a "natural" invisible *Ring* that causes the water vapor of vast air masses to condense. When these air masses slowly heat up and gain moisture over vast areas they produce clouds as symptoms that the dew point has been reached. When the densities of the water vapor and the air temperature are high, the cloud formation process can lead to the formation of large cumulus clouds, repeating on a large scale, with more marked dynamics, the control process that produces "small cumulus clouds." The air below the large cumulus clouds, pushed by the cold air, rises with impressive speed and force, dragging the lovers of gliding, caught unawares, to such high altitudes that the unfortunate victims suffer frostbite. While fair-weather cumulus form at relatively modest altitudes, where the air is only a few degrees colder, large cumulus, several kilometers high, encounter very low temperatures (the adiabatic vertical gradient corresponds to a fall of around 6.5 °C/km).

Finally, note that the formation of large cumulus clouds can arise from the coming together of masses of air with different densities and temperatures as a result of differences in atmospheric pressure. Since it is heavier, the shifting mass of cold air (or the cold front) normally wedges in below the lighter mass of hot air, thus producing a sudden cooling of the latter, which releases a large amount of vapor into the atmosphere. It is this vapor which, in fact, gives rise to large cumulus clouds that can even lead to thunderstorms.

The formation of fog at ground level derives from the same natural *Ring* that forms clouds in the atmosphere. When evaporation is uniform over a large area and there are no air bubbles with different densities, thus leading to a fall in air temperature toward the dew point, then humidity condenses at the ground level. Fog is nothing other than the symptom of the elimination of the difference between the effective temperature and the dew point; in other words, the symptom that the density has reached its saturation point, thus eliminating the density error and giving off water vapor. On cold winter days the same invisible *Ring* acts to form frost. Instead of falling to the dew point, the temperature must fall to the *frost point*, where the water vapor is transformed into ice. An entirely analogous phenomenon during cool summer evenings is the formation of dew.

6.5   Planetary *Rings*: "Gaia" and Daisyworld

If we broaden our observational horizon we can note that a large number of natural and invisible control systems are operating on our planet in a manner similar to that of the *Rings* we have examined in Sect. 6.4. These systems give rise to vast recurring phenomena at a planetary level; for example, the control system that regulates the water, or hydrologic cycle, the formation of permanent winds, trade winds, and monsoons, and the formation of ocean currents.

I shall examine the water cycle model in Sect. 6.10.1. Here I wish to look at the highly important system that controls the dynamics of *ocean currents*, which are in fact produced by the action of an *invisible Ring* that operates in a completely analogous manner as that which produces the fair-weather cumulus clouds in Fig. 6.9, though on a larger scale and in a different environment.

The oceans are crucial for the accumulation of heat and its distribution by means of the ocean currents:

> [The Oceans] possess a large heat capacity compared with the atmosphere, in other words a large quantity of heat is needed to raise the temperature of the oceans only slightly. In comparison, the entire heat capacity of the atmosphere is equivalent to less than three *metres* depth of water. That means that in a world that is warming, the oceans warm much more slowly than the atmosphere [...]. The oceans therefore exert a dominant control on the rate at which atmospheric changes occur. (Houghton 1994, p. 94)

Here I shall only consider how the *Ring* operates that produces and maintains over time the *Gulf Stream* in the northern hemisphere. The currents produced in the
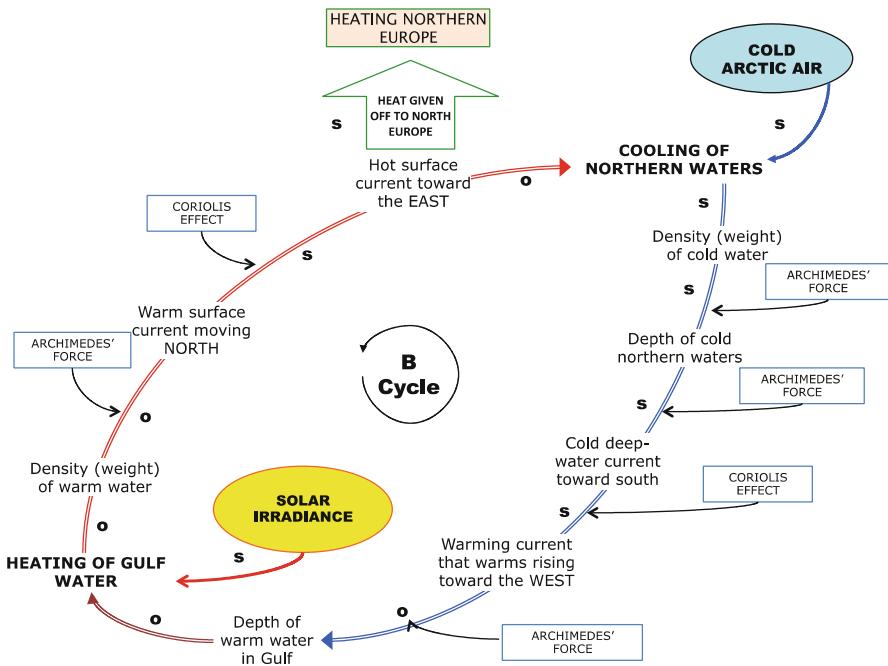
Fig. 6.10 The Gulf Stream *Ring*

southern hemisphere follow the same logic, though with dynamics that acts in a different direction and with variations that derive from this different orientation.

It is well known that the Gulf Stream originates in the Gulf of Mexico, follows the North American coast, crosses the Atlantic, and reaches Northern Europe, brushing the coasts of Portugal, Spain, France, and England, reaching as far as Scandinavia. From a physical point of view, the Gulf Stream is a giant flow of water kept distinct from the surrounding bodies of water due to its different temperature and saline density; it transports water at a rate of 74 million cubic meters per second at a speed of 1–2 m/s; in the middle Atlantic it is 400 km wide with an average temperature of 14 °C. From a climatic point of view it is a powerful means of transporting heat toward Europe, where it mitigates the climate by impeding the arrival of freezing arctic air.

The *Ring* in Fig. 6.10 illustrates how the Gulf Stream originates and is maintained. It is easy to see the similarity with the model in Fig. 6.9, with the obvious exceptions due to the fact it is a water current and not an air current, and that its direction of movement is determined not only by the heat–cold differential but also from the Coriolis effect that is a result of the earth's rotation (for a concise definition see http://science.yourdictionary.com/coriolis-effect).

The current flows continuously following a circular *logical* model—a cycle that repeats itself year after year—even if, in fact, its actual route is more irregular. Considering only the *logical system*, Fig. 6.10 clearly shows that the Gulf Stream is

kept active by two sources: a heat source (solar irradiation) located in the Gulf of Mexico, and a cold source (arctic air) from the North Pole. The heat source heats the water, which, becoming saltier and lighter due to Archimede's buoyancy force, remains at the surface, where the winds can begin to move it north.

The Coriolis effect from the earth's rotation directs the current toward the northeast, where the heated and salt-rich surface water begins to move north in the hemisphere, from the Central American coast toward the North European one. Proceeding northeast, the water encounters an area of colder air and starts to lose heat to the air (heating the air), thereby maintaining temperate the regions of Northern Europe. Once it has given off its heat, the increasingly colder water (due to the arctic climates of the northernmost regions) becomes even denser and heavier, so that the current must descend in depth due to Archimedes' force, subsequently moving southwest across the ocean as a result of the Coriolis effect. Moving down toward the equator in deep waters, the Gulf Stream runs into warmer water and regains sufficient heat to bring it to the surface in the Gulf of Mexico, where once again it is significantly heated before it again flows north, tracing a gigantic three-dimensional "8," after which the cycle starts again.

Now let us turn to other relevant "natural," invisible *Rings* that lead to the maintenance of acceptable temperatures on a planet populated by biological species that can alter temperatures simply through variables related to their numbers. Let us take as an example Daisyworld, an abstract world created by James Lovelock in his fundamental work, "Gaia" (written with the contribution of the biologist, Lynn Margulis). Lovelock sought to demonstrate how the earth, as a unitary system made up of a multitude of inseparably connected physical and biological variables, can in principle be considered an evolving "living being" that maintains itself in equilibrium over time through a holarchy of reinforcing and balancing loops that, despite the fact they are invisible, operate on a planetary scale.

> We now see that the air, the ocean and the soil are much more than a mere environment for life; they are a part of life itself. Thus the air is to life just as is the fur to a cat or the nest for a bird [. . .] There is nothing unusual in the idea of life on Earth interacting with the air, sea and rocks, but it took a view from outside to glimpse the possibility that this combination might consist of a single giant living system and one with the capacity to keep the Earth always at a state most favorable for the life upon it. An entity comprising a whole planet and with a powerful capacity to regulate the climate needs a name to match. It was the novelist William Golding who proposed the name Gaia. Gladly we accepted his suggestion and Gaia is also the name of the hypothesis of science which postulates that the climate and the composition of the Earth always are close to an optimum for whatever life inhabits it. (Lovelock 2011, Online)

Since one of the Earth's problems is the maintenance of the thermal equilibrium that allows various species to remain alive over time, Lovelock and Watson create Daisyworld, an imaginary ecosystem, to show that this equilibrium can be maintained—despite variations in insulation—thanks to the *Ring* based on the unconscious variation of biological masses.

> This book is the story of Gaia, about getting to know her without understanding what she is. Now twenty-six years on, I know her better and see that in this first book I made mistakes.

Some were serious, such as the idea that the Earth was kept comfortable by and for its inhabitants, the living organisms. I failed to make clear that it was not the biosphere alone that did the regulating but the whole thing, life, the air, the oceans, and the rocks. The entire surface of the Earth including life, is a self-regulating entity and this is what I mean by Gaia (Lovelock 1979, Preface).

Lynn Margulis, the coauthor of Gaia hypotheses, is more careful to avoid controversial figures of speech than is Lovelock. In 1979 she wrote, in particular, that only homeorhetic and not homeostatic balances are involved: that is, the composition of Earth's atmosphere, hydrosphere, and lithosphere are regulated around "set points" as in homeostasis, but those set points change with time. (Environment 2013, Online)

Daisyworld is viewed as a simplified hypothetical world on a planet that orbits around a star that is similar to the sun. Daisyworld is covered in daisies (imaginative component), which belong to two specimen populations of opposite color: the *white* Daisies, which flourish in hot climates, since the white color allows them to reflect most of the sun's light and thus not die from scorching, and the *black* Daisies, which instead have adapted to living in cold environments, since their dark color absorbs a good part of the solar light, thereby supplying them with the energy to live.

Watson and Lovelock (1983) show how this simplified ecosystem can self-regulate, regarding both the thermal equilibrium and the biological masses it is composed of, and indefinitely resist random limited variations in the heat emitted by the sun-like star.

When I first tried the Daisyworld model I was surprised and delighted by the strong regulation of planetary temperature that came from the simple competitive growth of plants with dark and light shades. I did not invent these models because I thought that daisies, or any other dark- and light-colored plants, regulate the Earth's temperature by changing the balance between the heat received from the Sun and that lost to space. (Lovelock 1988, p. 37)

I wish to demonstrate that the control system by which Daisyworld maintains itself in thermal equilibrium is composed of the joint action of two *Rings*, as shown in the Causal Loop Diagram in Fig. 6.11.

We must assume above all that there are vital temperature intervals for the two daisy populations. For the white daisies (WD), which live in the heat, there must be a minimum temperature, $C_{\text{MIN-WHITE}}$, below which there is no guarantee of the vital conditions for life, and above which (until a maximum temperature) they reproduce/die out at a *vital action rate*," $g(\text{WD}/C)$," for each degree of temperature (the rates of reproduction and extinction do not necessarily have to coincide and be constant). For the black daisies (BD), on the other hand, we assume there is a maximum temperature, $C_{\text{MAX-BLACK}}$, above which they cannot survive, but below which they reproduce/die out at a given *vital rate*, "$g(\text{BD}/C)$."

The mass of white and black daisies determines the reflecting surface according to the coefficients "m/r," which, if necessary, are separated into white and black daisies. Variations in the reflecting surface determine the variation in the temperature according to the coefficient "$h(C)$."

Let us agree that the system is temporarily in equilibrium at a temperature of $C^* = 40\,°\text{C}$, and let us assume that, due to some accidental phenomenon, there is
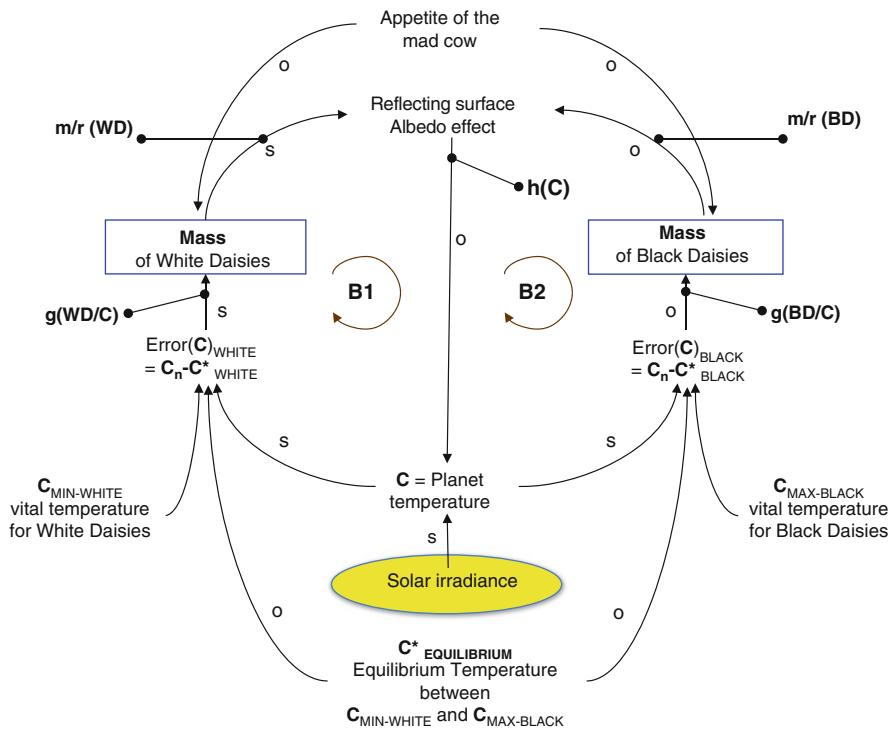
Fig. 6.11 Daisyworld and self-regulation (the model will be simulated in Sect. 6.10.2).

more sunlight, or that the sun, if possible, is hotter (a convenient hypothesis), for example $C_3 = 60\,^\circ$C at period $n = 3$ (see the simulation in Fig. 6.22 below).

What happens? Reading the model in Fig. 6.11 from left to right, we deduce that when the greater irradiance raises the planet's temperature, according to the link in direction "**s**," the white daisies have an error$(C)_{\text{WHITE}} = C_n - C^*_{\text{WHITE}} = 20\,^\circ$C with respect to the equilibrium temperature (note that, in order to make the description of the model more understandable, error(C) is calculated as the difference between $C_n - C^*_{\text{WHITE}}$). This positive error produces an increase in the mass of white daisies at the rate "$g(\text{WD}/C)$," since they are used to living in the heat. However, this causes an increase in the reflecting surface based on the coefficient "$m/r(\text{WD})$" (direction "**s**"), which reduces the planet's temperature according to the *action rate* "$h(C) < 0$" (direction "**o**"); this reduction lowers error$(C)_{\text{WHITE}}$ and causes a reduction in their mass, thereby reducing the reflecting power and initiating a new balancing cycle. This represents the first balancing loop, [**B1**] (there is only a single "**o**").

Operating simultaneously is the loop [**B2**], shown on the right side, which refers to the black daisies (BD). The increase in the planet's temperature increases error $(C)_{\text{BLACK}}$, which, because of the link in direction "**o**," reduces the mass of black daisies at the rate "$g(\text{BD}/C)$," since these are adapted to cold climates; the reduced

mass increases the reflecting surface (direction "**o**"), and this causes a reduction in the planet's temperature according to "$h(C)$." Thus loop [**B2**] is formed (there are two "**o**s") which, together with [**B1**], rebalances the temperature and the size of the biological masses (see the simulation in Sect. 6.10.2).

Obviously, if the temperature of the planet should rise above the maximum allowed for black daisies, $T_{\text{MAX-BLACK}}$, or fall below the minimum for white ones, $T_{\text{MIN-WHITE}}$, there would be a mass extinction of one of the populations; however, equilibrium could still be reached if the action and reaction rates of the surviving population permitted this by restoring the climate conditions that could favor the spontaneous regeneration of the extinct population.

Lovelock himself recognized the importance of identifying the survival limits of daisies.

> Very few assumptions are made in this model. It is not necessary to invoke foresight or planning by the daisies. It is merely assumed that the growth of daisies can affect planetary temperature, and vice versa. Note that the mechanism works equally well whatever the direction of the effects. Black daisies would have done as well. All that is required is that the albedo when the daisies are present be different from that of the bare ground. The assumption that the growth of daisies is restricted to a narrow range of temperatures is crucial to the working of the mechanism, but all mainstream life is observed to be limited within this same narrow range. (Lovelock 1988, p. 57)

The automatic natural *Ring* that regulates the temperature on Daisyworld strikingly resembles the two-lever temperature control system of Fig. 4.3, which controlled the water temperature of a shower with two separate taps. The mass of white daisies corresponds to the flow of cold water; the mass of black daisies to the flow of hot water.

However, there is an important difference: in Fig. 4.3 the temperature objective is set by the governor–manager and the control system regulates the flows from the hot and cold water taps so as to ensure the objective is reached. The Daisyworld model, on the other hand, is basically a self-regulating system for the two daisy populations rather than a control system for the planet's temperature, which is maintained through self-regulation.

We can make Daisyworld exactly like the shower system with two taps by choosing as the controlled variable the planet's temperature and as the objective the temperature that leads to equilibrium between the daisy masses. In reality, systems thinking teaches us that, since the three variables (white and black daisies, and temperature) are linked in a single dynamic system, none of the three rules is regulated by the others; instead all three are interconnected variables in a single system.

Another analogy is also possible: the Daisyworld model acts similarly to the control system that leads to the stabilization of prices through variations in demand and supply, shown in simplified form in Fig. 4.26. Though Daisyworld operates in a similar way, it differs from demand–supply equilibrium in its effects, since the objective of the latter control system is to maintain demand and supply equal through price variations. The Daisyworld model, on the other hand, has the objective of restoring the equilibrium temperature, which can also be achieved through different quantities of white and black daisies.

6.6   Control System for Global Warming

Once we have understood the operational mechanism of the Daisyworld model, a question naturally arises: does the model for temperature regulation in Daisyworld also apply to our planet? The answer appears to be no.

In fact, the Earth's temperature has continued to increase for decades, and it is difficult to imagine some form of automatic *Ring* that can return it to stable levels. Heating instead produces the opposite effect to that foreseen for Daisyworld, and the reason for this is easy to comprehend: the *albedo effect* which, in Daisyworld, was regulated by variations in the amounts of white daisies, is mainly produced on our planet by the extension of the polar ice cap. Moreover, on Earth the role of the black daisies is attributable to the amount of forests, woods and dark algae in the oceans. On Daisyworld an increase in temperature was automatically "balanced" by an increase in the amount of white daisies and a reduction in the amount of black daisies, leading to the forming of the two *balancing* loops shown in Fig. 6.11 and summarized in the left-hand side of the model in Fig. 6.12.

On our planet, on the other hand, an increase in temperature produces two *reinforcing* loops. As we can see on the right-hand side of Fig. 6.12, loop [**R1**] originates because the increase in temperature causes the surface area of the polar ice caps to shrink. Loop [**R2**] occurs because the increase in temperature makes the forests, woods and algae increasingly lush and extensive, thereby reducing the albedo effect rather than increasing it, as would be necessary to return temperature increases to an equilibrium state. The joint effect of the two reinforcing loops is to gradually increase global warming. Another difference between Daisyworld and Earth is that the latter is characterized by the *greenhouse effect*, to which human activity contributes. By amplifying the effects of solar radiation, the *global warming effect* is intensified.

The global warming phenomenon is produced by numerous causes, many of which are due to man's activities which, following the archetype of "short-term, local and individual preference" (Sect. 1.6.6 and Fig. 1.19), have increased
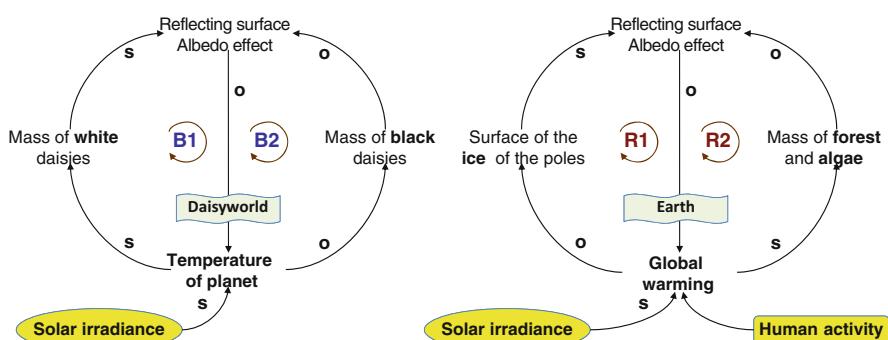
Fig. 6.12 Daisyworld and the Earth in comparison
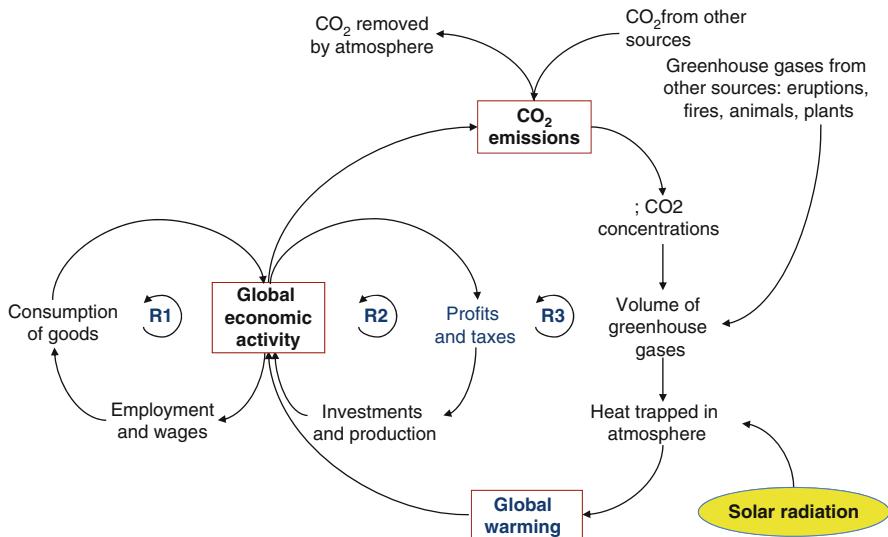
Fig. 6.13 A model for global warming

greenhouse gases in the atmosphere (carbon dioxide, methane, nitrous oxide, hydrofluorocarbons, perfluorocarbons, and sulfur hexafluoride).

> Three major pollution issues are often put together in people's minds: global warming, ozone depletion (the ozone hole) and acid rain. Although there are links between the science of these three issues (the chemicals which deplete ozone and the particles which arc involved in the formation of acid rain also contribute to global warming), they are essentially three distinct problems. Their most important common feature is their large scale. (Houghton 1994, Preface)

Among the many models available for understanding how global warming is produced, especially useful is the one by Peter Senge (1990, p. 345), which is presented in Fig. 6.13.

Senge's model is so clear there is no need for comment. The growth in global warming caused by loop [**R3**]—as a consequence of economic development as well as of the other sources of CO_2 emissions—seems unstoppable, and this produces and will continue to produce more and more damage due to unimaginable increases in sea levels, weather disturbances, the breakup of glaciers and drought, etc., in less time than predicted. For some (Maslin 2004) this damage—if we leave out population effects—could speed up economic development: houses to rebuild, entire cities to be moved, new products, etc., thereby increasing growth even more. Whatever its causes, this phenomenon can be regulated by a multi-lever control system (involving both artificial and natural levers), which, without claiming to be complete, is indicated in Fig. 6.14, which expands on Fig. 6.13.

The control system that regulates global warming uses three orders of levers, which are indicated in the lower lines in the figure. The operational and extraordinary levers are *artificial*; their action is specified in the dark boxes and their effects
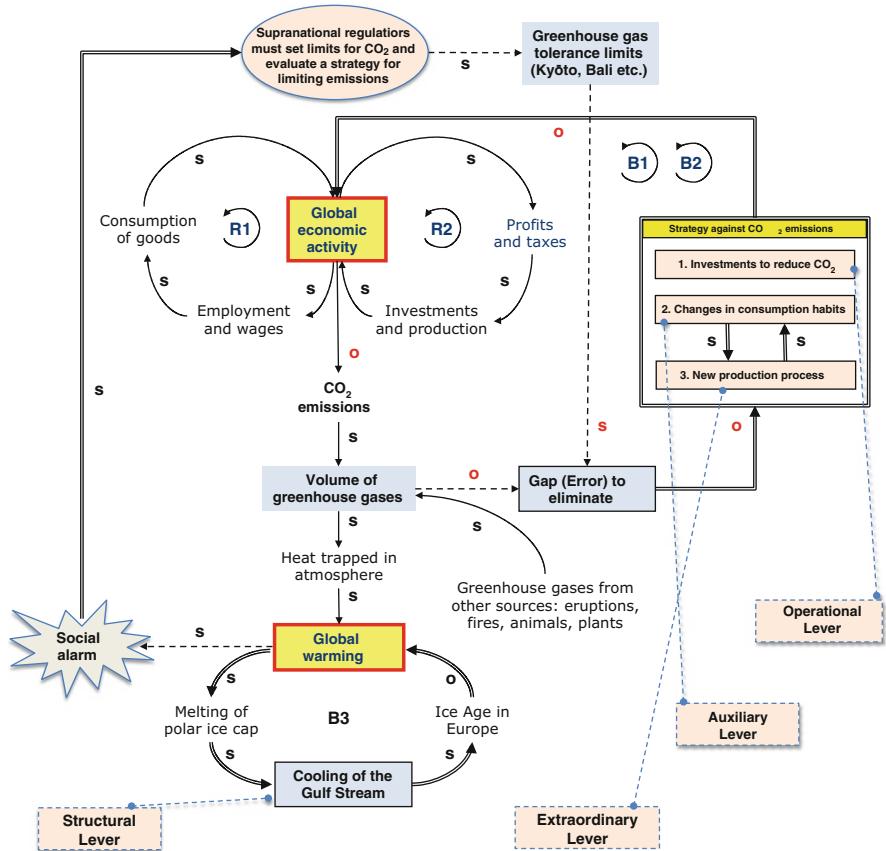
Fig. 6.14   Model of a control system for global warming (derived from Fig. 6.13)

represented by the arrows with the broken lines. The *natural* lever is structural in nature; it is shown in the lower right, and its action is represented by the double-lined arrows.

To orient ourselves in the challenging model in Fig. 6.14, I suggest the following reading sequence. When *global warming* is perceived and its effect begins to become worrisome, a *social alarm* arises (shaded box), which becomes ever more intense ("**s**"), producing a political movement that leads supranational regulators to set tolerance *limits to greenhouse gases* ("**o**") through a series of international agreements, among which the Kyoto protocol—signed in Kyoto on December 11, 1997, by over 160 countries at the COP3 Conference during the United Nations Framework Convention on Climate Change (UNFCCC)—along with other subsequent international agreements. These treaties oblige industrialized countries to reduce emissions of polluting substances (by at least 5.2 % from 2008 to 2012, with respect to emissions in 1990, which is taken as the base year). Recently the Bali agreement (2007) has established a drastic reduction (25–40 % by 2020) in greenhouse gases.

In order to reduce the gap between actual levels of CO_2 and those set by the regulators, the nations that sign the treaties must come up with a strategy composed of various levers (shown in the square on the right-hand side). These strategies activate above all the operational levers (box 1), which consist of *investments* to finance measures to reduce as much as possible CO_2 emissions, in order to bring them below the tolerance limits, thereby producing loop [**B1**] (these measures are too great in number to be detailed in the model). However, the social alarm also activates two other extraordinary levers: spreading a culture to *change consumption habits* (box 2), trying to limit the consumption of goods requiring a greater production of greenhouse gases, and the consequent change in *production processes* (box 3), again with the aim of reducing emissions of CO_2 and other greenhouse gases, thus producing loop [**B2**] in Fig. 6.14. Levers 2 and 3 are independent but mutually reinforcing.

The *structural* lever complements the operation of the artificial levers and acts to modify the structure of the geographic environment in which global warming produces its negative effects. The overall action of this lever can be viewed as an autonomous *natural* and automatic control system integrated into the system that produces global warming. Such a natural automatic *Ring*—which produces loop [**B3**] in Fig. 6.14—is identified by some (Lemley 2002, Online; NASA 2004, Online; Melchizedek 2004, Online) as the cooling of the Gulf Stream due to global warming, which, by melting the polar glaciers, dumps a mass of fresh cold water, which is also poor in salt, into the Atlantic Ocean, thereby cooling the ocean currents.

Figure 6.15 adds elements to the Gulf Stream model in Fig. 6.10 to demonstrate in more detail how the natural *Ring* [**B3**] in Fig. 6.14 operates to control global warming. The mass of cold, fresh water that spills into the Atlantic Ocean from massive ice cap melting ("**s1**") prematurely cools the flow of the Gulf Stream ("**s2**"), which, once cooled, precipitately sinks toward the ocean floor ("**s3**" and "**s4**") and turns toward the equator in a southwesterly direction.

The gradual melting of the polar ice caps, caused by the global warming, produces masses of sweet, cold water that pour into the Atlantic Ocean, diminishing its salinity. This phenomenon interrupts the current's "motor," pushing the Gulf Stream to the bottom and slowing its velocity. This precipitant sinking and slowing down of the Gulf Stream causes a smaller mass of the current's warm water to reach the North Atlantic ("**o1**"), and thus a smaller transfer of heat ("**s5**"), thereby turning the water colder again, with the risk of new glacier formation ("**o2**"). This hypothesis is supported by a large amount of data, and even the study of past weather phenomena appears to show that the last ice age, some 8,500 years ago, was caused in fact by the collapse of a glacier barrier in the Atlantic.

Glaciation in Northern Europe increases the *albedo effect* ("**s6**") because of two joint effects, similar to what we have witnessed in Daisyworld: on the one hand, the extension of the ice mass in Northern Europe (which corresponds to the increase in the mass of white daisies), and on the other the reduction of the green mass of northern forests (corresponding to the reduction in the mass of black daisies). These two effects reduce global warming ("**o3**"), at least temporarily, forming loop [**B1**].

These effects are summarized in Fig. 6.16. Loop [**B**] controls global warming by using the lever of the temperature in the northern regions, which depends on the
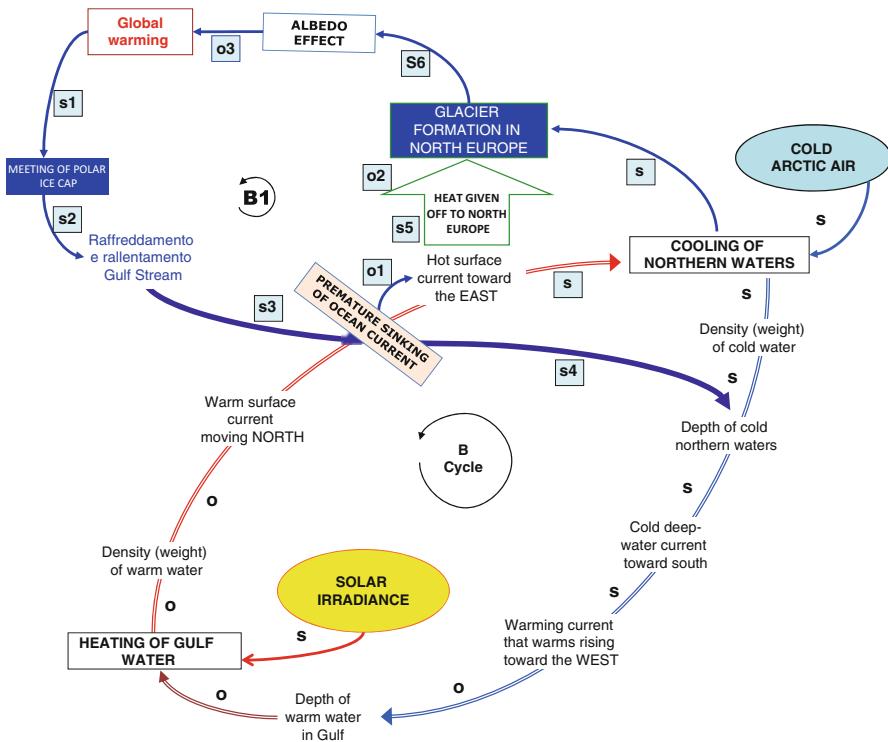
Fig. 6.15 The natural *Ring* to control global warming through the cooling of the Gulf Stream
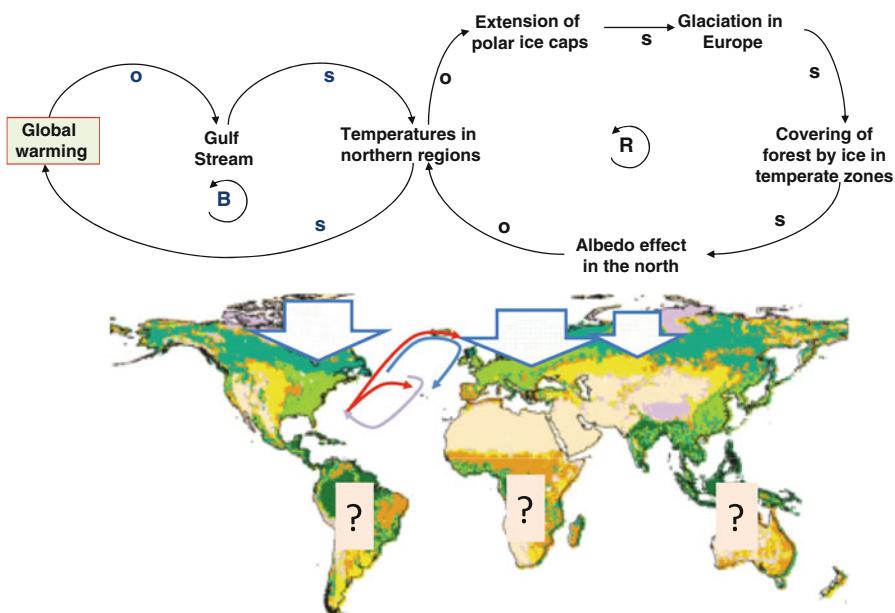
Fig. 6.16 The Daisyworld effect in the control of global warming

Gulf Stream; this lever functions as an effector to regulate this temperature, as illustrated in the model in Fig. 6.15. The sinking of the Gulf Stream reduces the temperature in the north, which activates loop [**R**], which in turn accentuates the albedo effect by regulating the two levers that extend the ice masses and cover the forest/green zone with ice. The natural *Ring* just described is, in fact, a *magic Ring*, which, if it ever spontaneously self-activates, is capable of dealing with the global warming problem.

Attempts have also been made to control global warming using other *artificial auxiliary levers*, but not all of those conceived of, even though apparently rational, are effective.

An artificial lever, which is simple to produce with the appropriate means, is that which produces the albedo effect on the oceans by spraying a wide area of clouds hovering above the oceans with seawater spray; the clouds then turn white, since they contain microscopic salt particles. In this way, the clouds will reflect a greater amount of solar radiation to produce the desired albedo effect, and as a result the earth will be heated up to a lesser extent.

> The idea behind the marine cloud-brightening (MCB) geoengineering technique is that seeding marine stratocumulus clouds with copious quantities of roughly monodisperse sub-micrometre sea water particles might significantly enhance the cloud droplet number concentration, and thereby the cloud albedo and possibly longevity. This would produce a cooling, which general circulation model (GCM) computations suggest could – subject to satisfactory resolution of technical and scientific problems identified herein – have the capacity to balance global warming up to the carbon dioxide-doubling point (Latham 2012, p. 4217).

> But at this point an unwanted surprise appears. Researchers have realized that the final effect was an undesired increase of rain on land, even to the point of causing terrible monsoons. Such violent phenomena occur, in fact, when the air over land is warmer than that over the oceans; the latter air then cools and produces clouds which are very white. (Caprara 2010, Online)

I should also reference a study of control systems for catastrophic events. Here I will just mention the studies to control great storms and tornadoes by bombarding them with appropriate chemical substances and those to control the route of asteroids and meteors if they should be on a collision course with Earth. The need for a control system to counter the collision of asteroid "2004 MN4," forecast for April 13, 2029, is quite urgent (Britt 2004; Hecht 2007).

Science fiction? Not any longer. When projects for control systems become reality, science fiction gives way to engineering (Fleming 2007).

6.7 *Rings* Acting on the Human Body

Now that we have observed the external world, it is appropriate to apply the discipline of control to our bodies, which, together with our minds, can be considered a psychophysical system that is kept alive by a network of natural and acquired control systems.

In fact, the same definition of life proposed by Herberto Maturana and Francisco Varela in their work, *Autopoiesis and Cognition* (1980) gains its operational validity from the existence of control systems as constituent structural elements of every living being. Living beings, man in particular, must be observed as closed structures due to their self-organization and conceived of as autopoietic and living machines or systems, that is, *homeostatic* systems which, thanks to a complex network of holonic multi-lever, multi-layer control systems, manage to maintain at the intercellular, intracellular, and organic levels a delicate equilibrium over time that allows them to continually reproduce the organization that defines them. These ideas are expressed most clearly by a direct quote from the work of Maturana and Varela:

> Autopoietic machines are homeostatic machines. Their peculiarity, however, does not lie in this but in the fundamental variable which they maintain constant. *An autopoietic machine is a machine organized (defined as a unity) as a network of processes of production (transformation and destruction) of components which: (i) through their interactions and transformations continuously regenerate and realize the network of processes (relations) that produced them; and (ii) constitute it (the machine) as a concrete unity in space, in which they (the components) exist by specifying the topological domain of its realization as such a network.* It follows that an autopoietic machine continuously generates and specifies its own organization through its operation as a system of production of its own components, and does this in the endless turnover of components under conditions of continuous perturbations and compensation of perturbations. Therefore, an autopoietic machine is homeostatic (or rather a relations-static) system which has its own organization (defining network of relations) as the fundamental variable which it maintains constant. (Maturana and Varela 1980, pp. 78–79)

This quote clearly shows that every *autopoietic machine* must regulate, or maintain constant, the *network of relations* that defines its own organization; in other words, without control systems that regulate the organizational variables, no system can be and remain autopoietic. Every autopoietic system thus has, in its *Rings* that regulate homeostasis, the vital condition for its existence.

On the basis of this definition the two authors also specify the concept of a living system:

> If living systems are machines, [then the fact] they are physical autopoietic machines is trivially obvious: they transform matter into themselves in a manner such that the product of their operation is their own organization. However, we deem the converse is also true: a physical system, if autopoietic, is living. In other words, we claim that the notion of autopoiesis is necessary and sufficient to characterize the organization of living systems. (Maturana and Varela 1980, p. 82)

As a result, the theory of autopoiesis—that is, the theory of living systems—clearly shows that our biological structure cannot maintain itself and develop over time without a network of *natural automatic Rings* that continually regenerates the network of vital processes by maintaining homeostasis and permitting autopoiesis to occur. Such *homeostatic natural automatic* control systems, together with those of *postural control*, which can also be nonautomatic, are fundamental for our existence. In the words of Norbert Wiener:

> Our homeostatic feedbacks have one general difference from our voluntary and our postural feedbacks: they tend to be slower. There are very few changes in physiological homeostasis – not even cerebral anemia – that produce serious or permanent damage in a small fraction of a second. (Wiener 1961, p. 115)

Not all relevant *Rings* that maintain homeostasis, even though in a disturbed environment, are known today, even if it is now clear that the levers used to regulate the various vital and cerebral parameters can also be chemical ones; biology and medicine are taking giant steps in uncovering the complex structure of microscopic and macroscopic *Rings*. The so-called *postural* control systems mentioned by Wiener are equally relevant in our daily lives, for us as well as for any animal characterized by articulation.

> In the human body, the motion of a hand or a finger involves a system with a large number of joints. The output is an additive vectorial combination of the outputs of all these joints. We have seen that, in general, a complex additive system like this cannot be stabilized by a single feedback. Correspondingly, the voluntary feedback by which we regulate the performance of a task through the observation of the amount by which it is not yet accomplished needs the backing up of other feedbacks. These we call postural feedbacks, and they are associated with the general maintenance of tone of the muscular system. (Wiener 1961, p. 107)

Postural control is usually achieved through a *multi-lever* control system—one which is entirely similar to that illustrated in Fig. 4.10 for focusing on an object on a horizontal plane—since it requires a number of levers to eliminate the distance, as Wiener recognizes in the above quote: "a complex additive system like this cannot be stabilized by a single feedback." However, I would also note that, in order to achieve the *postural control* of our joints, for example, those that allow us to grasp an object, the control system not only is *multi-lever* but also has features of a *holonic* control system just like the one presented in Sect. 3.8 for the control of the shifting of a point, P_t, in a multidimensional space in order to reach a point, P^*, which is set as the objective.

In fact, if we think about it, the action of moving our hand to grasp a small object (a pencil, pipe or cigarette, using Wiener's examples) in a fixed position is, from the control point of view, nothing other than the shifting of the present position of the hand, P_t, toward the position of the object, P^*. This is a movement in three-dimensional space, along with other, more precise spatial movements of the fingers, which must contract at the same time and at the "right point" in order to grasp the object. However, the postural control needed to grasp an object is complicated by the fact that the movement of the hand (and, on a smaller scale, the fingers) requires the joint action of various articulations—among which, the entire body, the bust, shoulder, arm, forearm, wrist, and, lastly, the fingers—each of which must, in turn, move in different spatial directions by means of the muscles that serve as translators, as shown in Fig. 6.17.

In fact, the control systems for each articulation must, though autonomous, be coordinated to permit the overall control of the movement of the hand toward the object–objective.

Operating similarly is the control of the movement of the two fingers that must "grab" the small object. The control system is further complicated if we consider the fact that, in order to move the articulations, it is necessary to control a number of muscles, which must be appropriately taught to guarantee "adequate" strength with
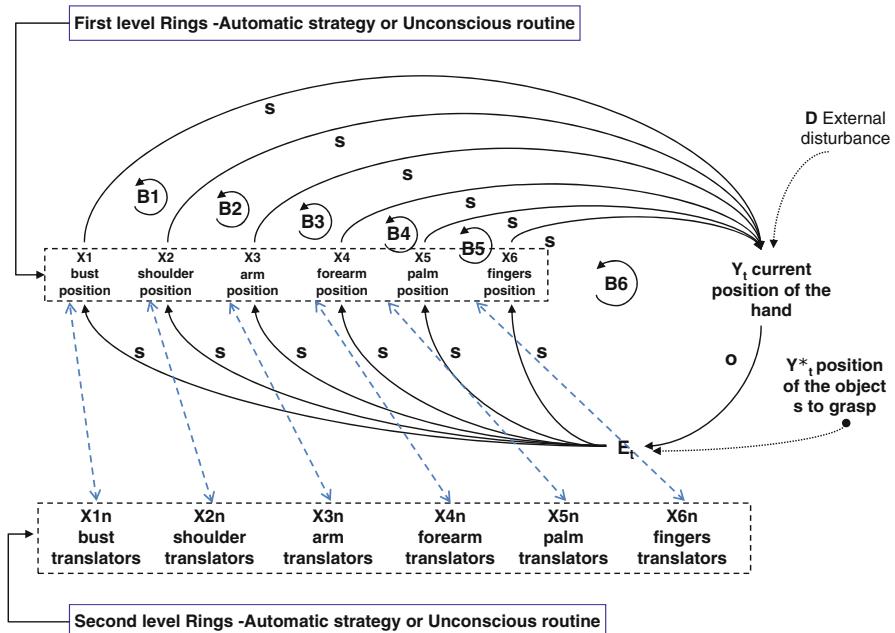
Fig. 6.17 Model of a holonic control system for grabbing a small object

respect to the size and weight of the object to grasp. Thus, a second level of translator control systems is formed, which is necessary for the functioning of the first-level control systems. The model in Fig. 6.17 becomes in all respects a *holonic* control system as soon as we realize that the control of the translators requires at least three third-level control systems, each of which is needed to modify the position of the articulation on a spatial coordinate, that is, to modify the tension of the muscles that change the direction of the movement of each articulation. All control systems, at the various levels, are coordinated by a precise automatic strategy, that is, by *unconscious routines* which in part are innate and in part derive from our experiences beginning in our childhood.

We could zoom in even more, but even without considering in detail the control of the fascia and individual muscular fibers, which make up the effectors, a conclusion becomes immediately clear: grasping a small object, no matter how simple it may seem to us, requires a large number of *Rings* which are coordinated and cooperate with one another. The operation of the "*macro*" global *Ring* that eliminates the gap between the position of the hand and object requires, in turn, the action of a number of "*meso*" *Rings*, which control the variation in the position of the individual articulations and, in an even more detailed manner, the action of the "*micro*" *Rings* that control the effector muscles of the other higher-level *Rings*.

Generalizing the above discussion, the control of the movement from a given point that requires a *series of articulations* can be carried out by a holonic, multi-lever, control system (which is probably also multi-objective) made up of control

systems of various levels. The further down we move to the sublevels of such systems, the more automatic and involuntary their operation becomes. As a result, such a holonic system must operate based on a strategy (and, when necessary, a policy as well) that is translated into *motor planes of action*, or in *physiological routines*—which we are thus not aware of—that regulate the levers in an *unconscious manner* to allow us to focus our *conscious* attention on the higher-level "macro" *Ring* needed to grasp the object. In this sense, we can clearly understand Wiener's words when he describes postural control systems with articulations:

> We have thus examples of negative feedbacks to stabilize temperature and negative feedbacks to stabilize velocity. There are also negative feedbacks to stabilize position, as in the case of the steering engines of a ship, which are actuated by the angular difference between the position of the wheel and the position of the rudder, and always act so as to bring the position of the rudder into accord with that of the wheel. The feedback of voluntary activity is of this nature. We do not will the motions of certain muscles, and indeed we generally do not know which muscles are to be moved to accomplish a given task; we will, say, to pick up a cigarette. Our motion is regulated by some measure of the amount by which it has not yet been accomplished. (Wiener 1961, p. 97)

No comment could be clearer, but one thing is certain: without the action of the many *Rings* that operate "magically" in a coordinated and cooperative manner unbeknownst to us, postural control would not be possible. To better understand these conclusions, let us consider several of the many *Rings* that involve our motor and perceptual apparatuses that are operating at every moment in our daily lives. These apparatuses continually become honed starting from the first days of life, and we come to take no notice of them because their regulation becomes so automatic, thanks to the *motor plans of action* and *physiological routines* that control our limbs and muscles (at various levels), allowing us to concentrate on the voluntary controls. However, we can become aware of them through several simple experiments.

Try this very simple one. Place several objects in front of you (e.g., a sheet of paper, a book, pen, glass, or other small object), then close your eyes and try to grab the objects, one after the other, with your right hand. I am sure you will all succeed in grabbing the objects, even with your eyes closed, from remembering their position. Now repeat the experiment, but try to note the involuntary postural control and understand how your hand adapts the position of the fingers in a different and specific way depending on the object you are about to grab. Without our being aware of it, every day the magic *Rings* that control our fingers act automatically to allow us to do all the movements for grabbing, squeezing, releasing, etc. Once again with our eyes closed, if we try to turn on the light with the wall switch we realize how the magic *Rings* that control the position of our arm, hand and fingers direct our index finger to the switch we have to turn on. This is the power of the *instrumental* automatic postural control for the control system for achieving the objectives we consciously set for ourselves.

A second, equally simple and meaningful, experiment could also be useful. Sit at a desk, or a level surface, in front of a mirror with a piece of paper and pencil. With your right hand (if you are not left-handed) print your name in block letters on the

paper, with letters 8–10 cm high, or draw a figure of your choice of similar size. Now try to trace the letters, but using your left hand. By observing the imprecision in your tracing, which covers the original writing or drawing with clear-cut errors, you will understand how the postural control system for articulation acts to allow us to control the position of the left hand in tracing over the original drawing. Proceeding with the experiment (using first the right hand and then the left one), trace again the letters (perhaps turning over the paper), but looking at the sheet of paper in the mirror and not directly. It is easy to see how difficult controlling the direction of the pencil becomes and how many errors occur in tracing the new lines over the original ones. The explanation for this is clear: an action so unusual as tracing lines on a page observed from a mirror impedes the action of the normal *physiological routines*, which are replaced by conscious regulations that are slowly modified. With patience, and trying the experiment several times, you will see that the errors decrease the more you repeat the movement; the hand will begin to move the pencil with greater security the more the *physiological routines* are readjusted.

This experiment teaches us many things; in particular, it confirms that under normal conditions the activities that require the use of our limbs are regulated by automatic control systems of which we are not aware. The second lesson is that such control systems have been routinized by "daily use"; we all remember how much effort it took in kindergarten or elementary school to correctly trace letters and form words. The third lesson is that the control systems for our movements and actions increase our precision through repetition, training and experience.

These evident facts are rich in consequences since, without entering into the complex "world" of learning, it lets us understand that this is in large part linked to the improvement in the efficiency of control systems through greater experience of the manager, or even from the consolidation of experience in *routines* and *standardized strategies* that make the regulation of the levers automatic and precise. To convince ourselves of this conclusion, we need only observe how children are taught to speak and move their limbs through the continual repetition of controls and the correction of the imprecise use of control levers; and to observe how much effort is required to teach them to remain erect and walk. Similarly, consider how activating the complex control system to remain vertical during a bus or subway trip, without resorting to hand straps or poles, has become for us an almost unconscious action. Also consider the effort a driving student must make to learn to guide a car or to park, control processes which, for experienced drivers, have become *routine*.

Finally, I must mention the very powerful system of multiple *Rings* that operate in a coordinated and cooperative manner to allow us to speak. This represents a form of control which differs from the postural one but which is no less important. The control systems for words must at the same time control the breathing in and out of air as well as its expulsion through the larynx and vocal chords, which must in turn be modulated to produce the desired sounds, phonemes, words, and sentences with the correct pronunciation (independently of meaning). Moreover, it is necessary to control the movements of the jaw and, above all, the tongue and lips.

These involuntary *Rings*, of which we have no awareness, are magic because they operate with very high precision and incredible speed. Learning to speak

begins when we are infants, and it continues through our entire life. It is easy to observe, on the one hand, the effort and patience of teachers in teaching young children the right position of their mouths and tongues in order to transform their stuttering into phonemes and words; and, on the other hand, the many attempts by young children to imitate the sounds pronounced by their teachers. Moreover, all these facets of speaking are well known to those who, having to learn another language, must apply the control systems for words to regulate their pronunciation.

6.8 Control Systems for Survival as Psychophysical Entities

Man, as a living being, not only is an autopoietic system but can also be considered (to an outside observer) as a conscious cognitive system, since, with internal organs for memory, computation, and evaluation (preferences), he is able to compare objects, calculate information, and construct representations in order to couple himself successfully to the environment and survive, even by modifying his own structure in line with the variations permitted by his genetic and operative program (Walsh 1995; Von Krogh and Roos 1996).

> (1) A cognitive system is a system whose organization defines a domain of interactions in which it can act with relevance to the maintenance of itself; the process of cognition is the actual (inductive) acting or behaving in this domain. Living systems are cognitive systems, and living as a process is a process of cognition. This statement is valid for all organisms, with and without a nervous system.
>
> (2) If a living system enters into a cognitive interaction, its internal state is changed in a manner relevant to its maintenance, and it enters into a new interaction without a loss of identity (Maturana and Varela 1980, p. 13).
>
> Observers know and create their environment through interactions with it. This interaction involves an explicit or implicit prediction about the environment". (Uribe 1981, p. 51)

The preceding *sections* sought to demonstrate how the interactions of man (and of a wide variety of animal species) with his environment is made possible by the ubiquitous presence and efficient action of innumerable *Rings*, voluntary or automatic, which are necessary to recognize and obtain all the objects needed to survive, or to flee from harmful phenomena, based on models of acquired knowledge. I do not wish to return to this discussion, but rather to consider in this *section* several fundamental control systems that operate every day at the *macro level* to allow us to survive, such as psychophysical systems, in particular the *Rings* that enable us to search for energetic and metabolic inputs in the environment held to be useful to maintain the network of vital processes.

We are all aware that under normal conditions we do not eat to the point of indigestion or drink until we are ready to explode. We do not run until we are ready to drop, stay awake for days on end, look directly at the sun, or stay where the temperature is too low or too high. Control systems watch over us with appropriate
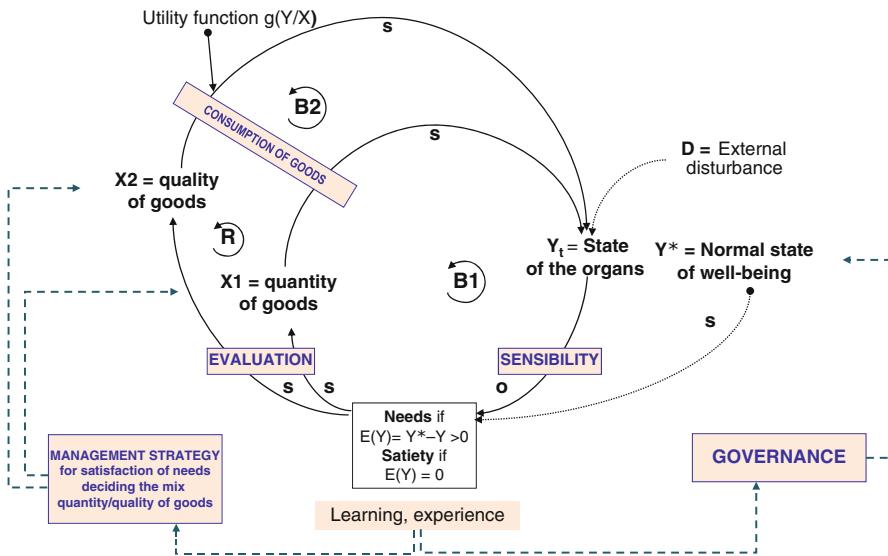
Fig. 6.18 General model of the control systems of needs

levers so that satiety, fatigue, blinding light, sleep, heat, and cold reveal the state of error with respect to the normal physiological values of the macro variables of our existence, and activate the corresponding levers to restore the conditions for our existence.

Let us begin by recognizing that, as systems possessing physical and mental activities, we are equipped with control systems that produce equilibrium in our psychophysical states. The general model of such systems is illustrated in Fig. 6.18. Such control systems, through the internal perceptive capacities of the governor–manager, detect at every instant the distance, or error, $E(Y)$, of the organism's state, Y_t, with respect to the "normal" states Y^*; that is, the breakdown of the psychophysical equilibrium. The error is perceived as negative, unpleasant, in need of being eliminated, minimized, or avoided. These unpleasant errors are called *needs* in the language of economics.

To eliminate needs we require *goods* of a certain *quality* and *quantity*. The quantity and quality of goods represents the two types of basic levers; the consumption of goods, quantitatively and qualitatively defined, modifies the state of our organism. When a need is perceived, man makes decisions to procure the necessary quantity and quality of the good to satisfy it. The *Ring* that allows us to deal with and manage our needs, shown in the general model in Fig. 6.18, must be supplemented by specific *Rings* for the various types of needs; these *Rings* are normally multi-lever and multi-objective control systems. We are cold (deviation) and the need arises to find clothes (first lever) to protect us and a house (second lever) for the night where we can light a fire (third lever); we are hot and thirsty (deviations), and so there is need for an ice-cold drink (first lever), a refreshing shower (second lever) and a

comfortable armchair in an air-conditioned room (third lever). We need to visit a friend who lives far away (objective), and in that moment we feel the need for transportation and choose to take our car (lever); we regularly feel the stimulus of hunger (deviation), and every time this occurs, the need for food arises (lever).

Needs are present in all aspects of our lives. It is useful to mention Maslow's (1943) attempt to classify needs. He proposed a scale, or pyramid of needs:

– *physiological* needs, represented by the need for the individual to maintain his psychophysical equilibrium;
– *safety* needs;
– *belongingness* and *love* needs, which operate at the social level;
– *esteem* needs, which correspond to our desire for self-esteem and social status; these come together to express the individual's need to define his identity, to distinguish himself in his environment, to acquire status and prestige;
– the need for *self-actualization*; this is at the top of the hierarchy of human needs and concerns man's need to realize his potential abilities, to develop and grow autonomously and continually, to be creative and full of ingenuity.

Fortunately man not only does aim to eliminate the *unpleasant states* that the many *Rings* psychologically and mentally reveal to him but also perceives *pleasant states* he wishes to acquire, maintain, or increase through appropriate control systems of pursuit, whose general model is shown in Fig. 6.19.

I shall call these states *aspirations*, which are produced by the gap, or error, between a desired state of well-being—which increases with the consumption of goods—and the state deriving from the consumption of available goods. The desire for a fashionable article of clothing does not derive so much from the *need* to protect ourselves from the cold as from the *aspiration* to be admired (objective); we feel the *need* for a car (more precisely, we feel the need for transportation) but we *aspire* to have a Ferrari or Lotus (objective), even if many other cars are able to satisfy the same need. But each time an *aspiration* is satisfied its level increases, so that it becomes more intense, thereby triggering a control system of pursuit toward an objective that increases the more we try to achieve it, based on the typical archetype of "Insatiability" (Fig. 5.12).

The action of control systems for *needs* and *aspirations* is so evident that their operation can be considered to be ingrained in our biological and social programming. With reference to Jacques Monod in his famous work, *Le hazard et la né cessité* (1970), we are conscious of the fact that, thanks to the existence of such control systems, man can be defined as a *teleonomic system* that tends, like all other animals, to ensure the survival of the species, of the group it identifies with based on the family's and individual's territory (tribe, nation, etc.).

All artifacts are the product of the activity of a living being that expresses in this way, and in a particularly evident manner, one of the fundamental features that characterizes all living beings, without exception: that of being an object endowed with a project which is represented within their structures and is carried out by means of their performance (for example, the creation of artifacts).

Rather than refusing to accept this idea (as some biologists have tried to do), it is instead indispensable to recognize it as being essential for the definition of living beings
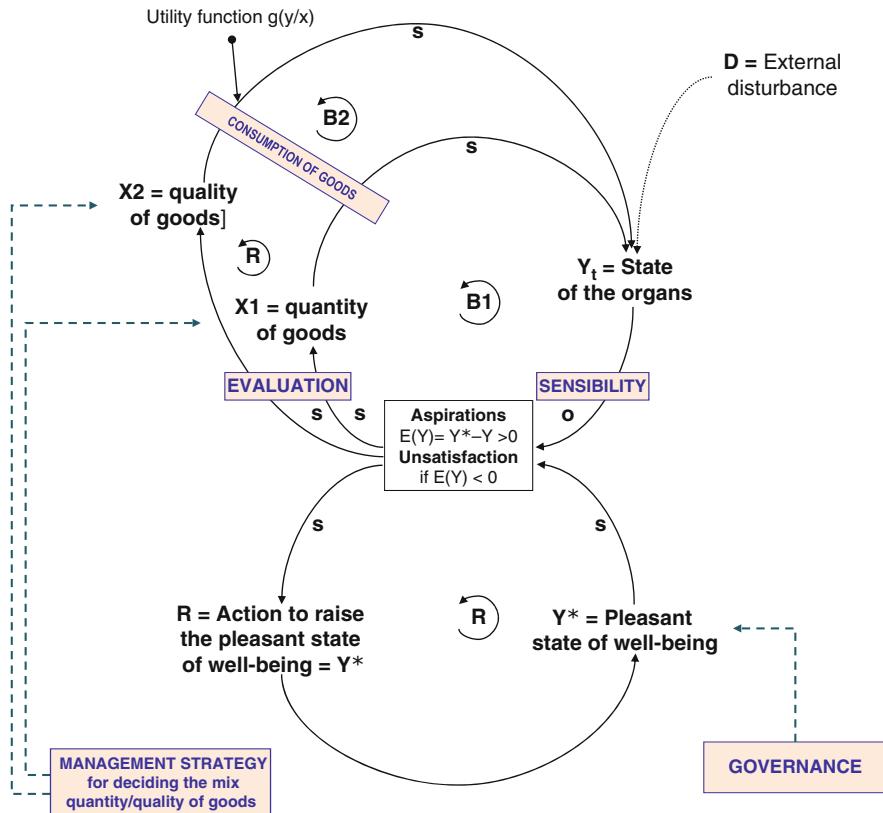
Fig. 6.19 General model of control systems for aspirations

themselves. We can say that the latter are distinguished from all other structures of systems in the universe by this property that we will denote as teleonomy.

We arbitrarily choose to define the essential teleonomic project as consisting in the transmission, from one generation to the next, of the unvarying contents characteristic of the species. All the structures, all the performances, all the activities which contribute to the success of the essential project will thus be called "teleonomic". (Monod 1970, pp. 22–25)[1]

---

[1] Tout artefact est un produit de l'activité d'un être vivant qui exprime ainsi, et de façon particulièrement évidente, l'une des propriétés fondamentales qui caractérisent tous les êtres vivants sans exception: celle d'être des objets doués d'un projet qu'à la fois ils représentent dans leurs structures et accomplissent par leurs performances (telles que, par exemple, la création d'artefacts). Plutôt que de refuser cette notion (ainsi que certains biologistes ont tenté de le faire), il est au contraire indispensable de le riconnaître comme essentielle à la définition même des tre vivants. Nuos dirons que ceux-ci se distinguent de tiutes les autres structures de tous systèmes présents dans l'univers, par cette propriété que nous appellerons la téléonomie) (Monod 1970, p. 22) [...] Nous choisissons arbitrairement de définir le projet tèléonomique essentiel comme consistant dans la trasmission, d'une génération à l'autre, du contenu d'invariance caractéristique de l'espèce. Toutes les structure, toutes les performances, toutes les activitès qui contribuent au seccès du projet essentiel seront donc dites "téléonomiques" (Monod, ibidem).

Teleonomy is achieved on the condition that the *Rings* that reveal the continual creation of *needs* and *aspirations* are active and that these lead to the search for levers to achieve the maximum satisfaction to ensure the maximum quality of life for the individual's survival, which is necessary to perpetuate the species to which it belongs. For this reason, in my opinion there is a close relationship between *teleonomy* and *autopoiesis*, since teleonomy—understood as the attitude of a species, as a collectivity, to perpetuate itself—can be considered the phenomenology that corresponds to autopoiesis—understood as the self-production of species with respect to the individuals.

> In effect teleonomy is teleology made respectable by Darwin, but generations of biologists have been schooled to avoid 'teleology' as if it were an incorrect construction in Latin grammar, and many feel more comfortable with a euphemism. (Dawkins 1982, cited by Barrows 2001, p. 705)

In fact, our way of life depends on the number and types of *needs* and *aspirations* we can satisfy, and above all on the way we achieve this satisfaction. These general comments make it easier to investigate and recognize the various *Rings* that maintain teleonomic conditions for the individual. Among the most evident are the control systems for food, water, fuel, electricity, and heating needs, those to regulate our health and hygienic conditions, to control the social relations of education, and to form thoughts about the self and others. The building of hospitals, hotels, vacation resorts, the continual increase in the size and efficiency of new firms, and the increase in highly gratifying jobs are all examples of action levers used in these control systems for needs and aspirations. However, we can easily see—as we have observed on several occasions above—that even our daily behavior is the result of the action of control systems that allow us to achieve our objectives in life. The action of *Rings* on our behavior is pervasive, and our behavioral life depends on these systems.

This should be clear, but it is useful nevertheless to reinforce this by presenting the interpretation of behavioral control provided by Perceptual Control Theory (PCT), a psychological-cognitive approach that bases the study of our behavior on the control systems that determine it. PCT focus on man—and more generally on cognitive systems (entities or agents)—trying to explain how such entities control what happens to them by investigating the relation between actions and objectives, perceptions and actions, and perceptions and reality based on the typical logic of control systems (Carver and Scheier 1981).

> Perceptual Control Theory (PCT) is a theory of how 'control' works to the level of detail that control can be modelled with precision. It regards life as a process of control and proposes that all living things are purposeful in doing what they do.
>
>   To put it simply, people need a goal, a means, a resource and they need to pay attention to the results – which are the effects on our own experiences. These elements weave together in a closed circle that builds control. The goal is what the person can hold in mind and strive towards – to be a fireman, to be loyal, to be a good son or daughter. In PCT this is known as a reference value. The means are the behaviours, skills, work, emotional attunement to others and ways of thinking (outputs) that a person exercises to achieve this goal. The resource is the feature of the environment, including people around us, who can help us achieve this goal. Paying attention to our own experiences closes the loop because

we perceive our results and compare them to our goal. If we have not reached our goal yet, we continue to develop the means to achieve it. When this cycle is allowed to prosper at a young age, a sense of purpose is built up that brings in the local environment and one's own community as part of one's own sense of purpose rather than as a force to be challenged and attacked. (Mansell 2011, Online)

The guiding principle of PCT is that human behavior is not the output of a system of transformation of *stimuli* into *actions*—as in traditional behavioral theory—but the input (X) of a control system to achieve (Y) what man wants: his objectives, desires, and aspirations, (Y^*). In other words, the stimulus to action is considered the perception of an error, $E(Y)$, between the present position (Y) of the agent and the objective (Y^*) he wishes to achieve. In the Web site of the Control Systems Group, this logic is summarized as follows:

> PCT is a theory about 'control' in living things, like people, cats and fish. It's about the control of what they perceive – what they notice, look at, feel, hear, taste and smell.
>
> The living thing is in the driver's seat. PCT is a theory from the person's own point of view – or the animal's own point of view...
>
> When things are working normally, the person gets to experience what they want to experience. It is 'just right' – like the perfect cup of coffee or tea...
>
> But how do they know this is happening?
>
> According to PCT, the person compares a 'standard' – what they want – with what they are experiencing right now – their perception.
>
> The difference between the two – the discrepancy or error is being measured. The bigger the error the more the effort the person makes to reduce it, until the error is zero – this means they get what they want. (http://www.pctweb.org/whatis/whatispct.html)

I wanted to include this long citation because it clearly confirms that our behavior is nothing other than a continuous chain of actions to activate the control levers necessary to achieve an objective. Thus, according to PCT, our actions are the physical activities of the control system levers (X) that are activated based on the mental processes for detecting error (attention) and for regulating behavior aimed at the objective. If we add that the objective (Y^*) is the output (Y) to achieve (grab a pen, take a trip to a certain place, take the bus, etc.) and is generated by the same mental process that produces the actions to undertake (X) to activate the levers (output to carry out), then the following conclusion is valid: every human action must be interpreted in the context of the control systems of which it is a part. This produces the following *corollary*: man is a *teleological system*. He does not act based on "cause → effects," "actions → reactions," "stimuli → responses," "disturbances → adaptations," but to achieve *objectives* of various kinds through *actions* and *behaviors* that represent the levers (X)—or inputs—of the control system that aims at eliminating the error perceived between the present position of the individual (Y) and the objective, (Y^*) in accordance with the general model of a control system.

William Powers, in his book *Behavior: The Control of Perception* (1973), proposed a more complex formulation of PCT. He starts from the general thesis that human behavior is activated by a control system to eliminate the effects of the external disturbances [Dm] on the M vital variables [Ym] (where m varies from 1 to M); these disturbances produce the errors $E(Ym)$, which we perceive as the stimuli

that activate our behavior (Xn). However, Powers observes that these perceptions can be divided into a hierarchical order covering nine levels: (1) Intensity, (2) Sensation or Vector, (3) Configuration, (4) Transitions, (5) Sequence, (6) Relationship, (7) Program, (8) Principles, and (9) System concepts. This order is not obligatory but rather only represents a proposal for further refinement; in fact, Powers subsequently introduced two new levels: 10 events and 11 categories, thereby also modifying the hierarchical order. What is important to note is that the perceptions at each level activate not only the behavior needed to control those perceptions but also the behavior to control the perceptions at the higher lever. Powers therefore recognizes that human behavior must be viewed as a typical holarchy of perceptions and behavior at various levels. To emphasize the idea of a hierarchy of perceptions and hierarchical controls by means of coordinated behavior, Powers defined his new formulation as the Hierarchical Perceptual Control Theory, or HPCT, precisely because the theory includes a hierarchy of perceptions and a parallel hierarchy of control. I shall not go beyond this brief description of Powers' work; however, I would underscore that he illustrated the role of memory and imagination as means for reorganizing perceptions. He introduced a second-level control defined as the *reorganizing system*, which operates to change the organization of the perceptual control hierarchy (Y) so as to remove potentially harmful conflicts, $E(Y)$, and improve (X) the control of perceptual signals at all levels.

If we consider all the observations from this *section*, we see that the Discipline of Control has clearly demonstrated that man, as a psychophysical system, can exist and continue living only to the extent he is equipped with a holonic network of *Rings* that maintains autopoiesis and permits him to behave in a teleonomic and teleological manner. Even Ludwig von Bertalanffy, considered the founder of the general theory of systems, clearly recognized that:

> So a great variety of systems in technology and in living nature follow the feedback scheme, and it is well-known that a new discipline, called Cybernetics, was introduced by Norbert Wiener to deal with these phenomena. The theory tries to show that mechanisms of a feedback nature are the base of teleological or purposeful behavior in man-made machines as well as in living organisms, and in social systems. (von Bertalanffy 1968, p. 44)

Nevertheless, teleological behavior cannot be viewed as limited to activities needed for survival alone but must be extended to all those creative and artistic forms of behavior that denote man's innate tendency to achieve objectives involving his self-fulfillment and the improvement of his conditions in life. In order to fully understand how fundamental the above-mentioned action of the *Rings* are in guiding man's actions toward objectives with a strong *creative* intellectual content, we need only reflect on how such systems act to guide artists, painters, sculptors, musicians, architects, poets, etc., in producing their works.

Consider the number and type of magic *Rings* that have influenced the sensibility to sound, color, proportion and shape, and that have guided the movement of the eyes and limbs to allow Leonardo to paint the Mona Lisa, Michelangelo to sculpt the Pietá, Bernini to build his cathedrals, Beethoven to compose his symphonies, and Dante to write the Divine Comedy. If the control systems that have guided these

artists seem powerful, the cathedrals in Siena and Orvieto, with the splendor of their frescos, statues, paintings, inlaid marble and wood, stain-glassed windows and mosaics represent what is perhaps an unequalled sample of the joint action of billions of *Rings* that magically have guided the work of thousands of artists (and workers) toward the objectives conceived by their imagination. Every museum, architectural work, opera house, or library testifies to the incredible action of the magic *Rings* in leading man toward artistic creations of all kinds for his self-fulfillment.

6.9   *Rings* That Regulate Biological Clocks

To conclude my observations I would note that, like all living beings, our body has a number of "time" control systems that regulate most of our physiological functions at all levels. Our behavior, like that of the organisms of all living species, is conditioned by the variations in the rhythms of the sun and moon—during the year, the months, and the diurnal and nocturnal paths of the sun—as if such variations were controlled, with respect to given periods or instants memorized as objectives, by innate control systems, which are commonly called *biological clocks* and studied in chronobiology (Dunlap et al. 2003).

How do organisms "know" when it is time for a particular activity, say plants to detach the fruit from the branches; the chrysalis to emerge from its cocoon; penguins, other birds, and fish to migrate to lay their eggs? Why does a woman's pregnancy last 40 weeks while for other species the time period is different? Why do we feel the urge to eat, sleep, and wake up after an average number of hours?

The most intuitive answer is that all these cyclical events are regulated by biological clocks (Koukkari and Sothern 2006). In fact, we can assume that a *biological clock* is an *impulse* control system (Sect. 4.3) whose objective, Y^*, is the value of some type of *accumulation* variable Y, positive or negative, or the progressive elimination of some type of *dissipative* variable, as shown in the model in Fig. 6.20.

The gradual accumulation of values for Y continues until Y^* is reached. The gradual elimination of the dissipative variable continues until Y attains the value $Y^* = 0$. The $E(Y) = 0$ represents the signal that the physiological processes, until that moment inhibited, have been activated. The nature of the variable Y, whose values are accumulated or eliminated, depends on the type of biological clock and on the specific organism in which it operates. Chemical substances as well as frequency signals can be accumulated produced by the action variable X, for example, melatonin from the pineal glands, cortisol from the suprarenal glands, the circadian rhythm of intracellular potassium, and so on. In any event, it is clear that the term "biological clock" is used for convenience sake; however, there is no independent clock that "detects" the passing of time. Instead, there is a control
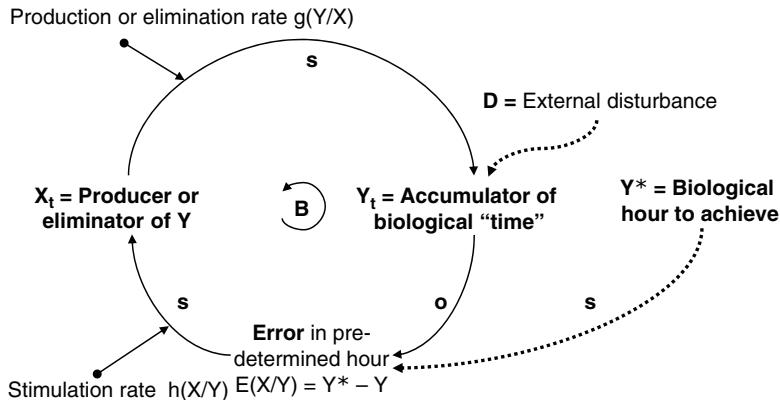
Fig. 6.20   The biological clock as an impulse control system

system that "detects" the accumulation or the elimination of some flow variables, usually chemical, biochemical or physical, until they reach levels (upper or lower) contained in the genetic programming of the biological individual (Fig. 6.20).

6.10   Complementary Material

6.10.1   The Water Cycle

Very briefly, the *water cycle* describes when water at ground level partially evaporates and is transferred to the atmosphere to form cloud masses which subsequently descend back to the ground in the form of rain, hail, or snow. The water cycle can be illustrated by the model in Fig. 6.21 (which does not claim to be complete) which, representing a continual, repetitive (though not closed) process, can in all respects be likened to an invisible planetary *Ring*.

It is well known that evaporation is a continuous phenomenon when the temperature is above zero, and that it is favored by air pressure and wind. Evaporation becomes quite intense as a result of solar irradiation over large areas, lands, forests and oceans. Unlike what occurs in small cumulus clouds that form due to convective movement, water vapor rises slowly, invisibly, increasing the density of large masses of air that can be moved by the wind. Clouds form due to water vapor condensation when the air masses encounter colder masses with different densities. If the temperature falls below the saturation temperature, the water vapor of the clouds generally *precipitates* in the form of rain; if the temperature falls further, this *precipitation* is in the form of snow.
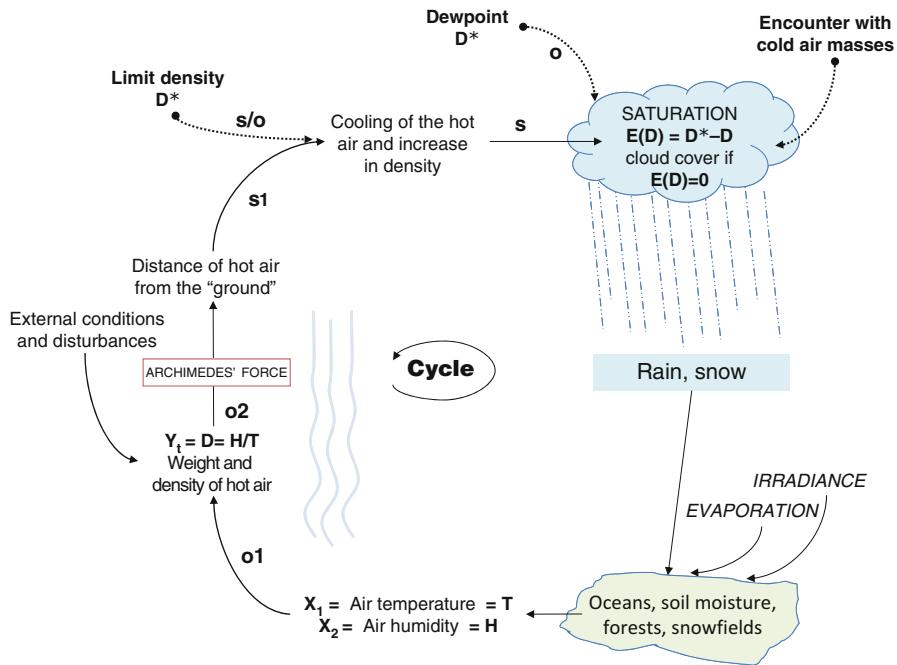
Fig. 6.21 The water cycle model

6.10.2 Daisyworld Dynamics

Watson and Lovelock (1983) produced a model simulating Daisyworld, assigning appropriate values for the growth of daisies based on temperature variations, translating the variation in the masses of white and black daisies into variations in the albedo effect, as indicated in the model in Fig. 6.11. Usually the decision to use only white and black daisies is based on the simplifying assumptions that the other daisy color gives no direct reproductive advantage and that the two extreme colors, black and white, make the feedbacks that regulate the entire system more effective. However, Lovelock also designed a model which includes grey daisies in order to make variations in planetary temperatures less drastic.

> With rare exceptions, biologists either ignored the models or remained as skeptical as ever. A persistent criticism from biologists was that, in a real world, daisies would have to use some of their energy to make pigment and therefore would be at a disadvantage compared with unpigmented gray daisies. In such a world no temperature regulation would take place. As they put it, "the gray daisies would cheat."
>
> Stimulated by their criticism I made a model with three species of daisies. All that the new model required was another set of equations to describe the temperature and growth of the gray daisy species. It was a matter of introducing sober-suited middle-management daisies to a world of colorful eccentrics. (Lovelock 1988, p. 45)

In order to take into account the three daisy populations, each sensitive in a different way to temperature variations, the model in Fig. 6.11 would have had to have been extended by a third loop and have behaved similarly to The Law of Dynamic Instability (Sect. 1.6.3).

Even if there are a large number of simulation applets for Daisyworld, which the reader can easily find on the Internet, I feel it useful to undertake an Excel simulation, which in my view provides the immediate advantage of allowing us to grasp the dynamics of the daisy populations and of the temperature under differing hypotheses. It is sufficient to construct the model based only on two populations of white and black daisies. Figure 6.22 shows the dynamics of the white and black daisy populations and of the temperatures in Daisyworld.

The action rates are indicated in the initial control panel, which assumes an equilibrium between the daisy masses at a temperature of $C^* = 40\,°\mathrm{C}$. At instant $n = 3$ there is a rise in the temperature of the C_3 disturbance to $20\,°\mathrm{C}$. The second column clearly shows that starting from instant $n = 4$ the number of white daisies begins to increase at the rate of $g(\mathrm{WD}/C) = 1\,\%$ (as if a flow of cold water were triggered), while at the same time the number of black daisies begins to decline at the rate $g(\mathrm{BD}/C) = 1\,\%$ (as if a flow of hot water were shut off), which, for simplicity's sake, coincides with that of the white daisies. The joint dynamics of the two daisy populations increases the albedo effect with a consequent gradual decline in the planet's temperature. Beginning at instant $n = 13$ the dynamics stabilize around the value WDmass $= 1,105$ and BDmass $= 904$, and the temperature in Daisyworld returns to $C_{13} = 40\,°\mathrm{C}$. A new equilibrium is achieved with different masses from the initial ones.

Let us further assume that at $n = 15$ a cold front occurs that lowers the temperature of the C_{15} disturbance to $-20\,°\mathrm{C}$. This lowering of the temperature in Daisyworld reduces the mass of white daisies, while that of the black daisies begins to increase, thereby producing a negative albedo effect that leads to an increase in the planet's temperature, which at $n = 25$ returns to $C_{25} = 40\,°\mathrm{C}$. Note that there is a change in the masses of white and black daisies with respect to the previous equilibrium attained at $n = 13$: the masses return to the initial level of WDmass $= 1,000$ and BDmass $= 1,000$, which includes a rounding off by one unit owing to the tolerance of 1 (control panel). The dynamics of the daisy masses and of the temperatures are shown in the graph in the lower part of Fig. 6.22.

Figure 6.23 presents a second simulation that more realistically provides for gradual cold impulses lasting 4 years, from $n = 4$ to $n = 7$, followed, after some time, by gradual cold impulses, from $n = 18$ to $n = 21$. From year $n = 10$ until $n = 13$ a random disturbance is introduced that strikes the white daisies, which are reduced in number due to the so-called "mad cow" effect, that is, a random reduction in the mass of daisies, as if they were eaten by a cow that randomly eats white daisies and black daisies.

As the model predicts, both the effects of the warm and cold impulses and the disturbances caused by the "mad cow" effect on the temperatures are always eliminated by the dynamics of the masses of the two daisy populations, as can clearly be seen in the graph at the bottom of Fig. 6.23.

W_0	1000	B_0	1000	h(C)	-0.1
g(WD/C)	1.0%	g(BD/C)	1.0%	C_0 =	40
m/r WD	1	m/r BD	1	C* =	40
				tolerance	1

n	WHITE Daisies mass	BLACK Daisies mass	DELTA albedo WL	DELTA albedo BL	Total albedo	C_n = Daisyworld temperature	C* OBJECTIVE	D_n DISTURBANCES	ERROR C_n – C*	mad cow WL	mad cow BL
0	1000	1000	0.0	0.0	0.0	40.0	40.0		0.0		
1	1000	1000	0.0	0.0	0.0	40.0	40.0		0.0		
2	1000	1000	0.0	0.0	0.0	40.0	40.0		0.0		
3	1000	1000	0.0	0.0	0.0	60.0	40.0	20.0	20.0		
4	1010	990	10.0	10.0	20.0	58.0	40.0		18.0		
5	1020	980	10.1	9.9	20.0	56.0	40.0		16.0		
6	1030	970	10.2	9.8	20.0	54.0	40.0		14.0		
7	1041	961	10.3	9.7	20.0	52.0	40.0		12.0		
8	1051	951	10.4	9.6	20.0	50.0	40.0		10.0		
9	1062	941	10.5	9.5	20.0	48.0	40.0		8.0		
10	1072	932	10.6	9.4	20.0	46.0	40.0		6.0		
11	1083	923	10.7	9.3	20.0	44.0	40.0		4.0		
12	1094	914	10.8	9.2	20.1	42.0	40.0		2.0		
13	1105	904	10.9	9.1	20.1	40.0	40.0		0.0		
14	1105	904	0.0	0.0	0.0	40.0	40.0		0.0		
15	1105	904	0.0	0.0	0.0	20.0	40.0	-20.0	-20.0		
16	1094	913	-11.0	-9.0	-20.1	22.0	40.0		-18.0		
17	1083	923	-10.9	-9.1	-20.1	24.0	40.0		-16.0		
18	1072	932	-10.8	-9.2	-20.1	26.0	40.0		-14.0		
19	1061	941	-10.7	-9.3	-20.0	28.0	40.0		-12.0		
20	1050	951	-10.6	-9.4	-20.0	30.0	40.0		-10.0		
21	1040	960	-10.5	-9.5	-20.0	32.0	40.0		-8.0		
22	1030	970	-10.4	-9.6	-20.0	34.0	40.0		-6.0		
23	1019	979	-10.3	-9.7	-20.0	36.0	40.0		-4.0		
24	1009	989	-10.2	-9.8	-20.0	38.0	40.0		-2.0		
25	999	999	-10.1	-9.9	-20.0	40.0	40.0		0.0		
26	999	999	0.0	0.0	0.0	40.0	40.0		0.0		
27	999	999	0.0	0.0	0.0	40.0	40.0		0.0		
28	999	999	0.0	0.0	0.0	40.0	40.0		0.0		
29	999	999	0.0	0.0	0.0	40.0	40.0		0.0		
30	999	999	0.0	0.0	0.0	40.0	40.0		0.0		

Temperature Dynamics

Populations Dynamics

White daisies mass
Black daisies mass

Fig. 6.22 First Daisyworld simulation

W$_0$	1000	B$_0$	1000	h(C)	−0.1
g(WD/C)	1.0%	g(BD/C)	1.0%	C$_0$ =	40
m/r WD	1	m/r BD	1	C* =	40
				tolerance	1

n	WHITE Daisies mass	BLACK Daisies mass	DELTA albedo WL	DELTA albedo BL	Total albedo	C$_n$ = Daisyworld temperature	C* OBJECTIVE	D$_n$ DISTURBANCES	ERROR C$_n$ − C*	mad cow WL	mad cow BL
0	1000	1000	0.0	0.0	0.0	40.0	40.0		0.0		
1	1000	1000	0.0	0.0	0.0	40.0	40.0		0.0		
2	1000	1000	0.0	0.0	0.0	40.0	40.0		0.0		
3	1000	1000	0.0	0.0	0.0	40.0	40.0		0.0		
4	1000	1000	0.0	0.0	0.0	42.0	40.0	2.0	2.0		
5	1010	990	10.0	10.0	20.0	44.0	40.0	4.0	4.0		
6	1020	980	10.1	9.9	20.0	48.0	40.0	6.0	8.0		
7	1030	970	10.2	9.8	20.0	54.0	40.0	8.0	14.0		
8	1041	961	10.3	9.7	20.0	52.0	40.0		12.0		
9	1051	951	10.4	9.6	20.0	50.0	40.0		10.0		
10	1052	941	0.5	9.5	10.0	49.0	40.0		9.0	−10.0	
11	1052	932	0.5	9.4	9.9	48.0	40.0		8.0	−10.0	
12	1053	923	0.5	9.3	9.8	47.0	40.0		7.0	−10.0	
13	1053	914	0.5	9.2	9.8	46.0	40.0		6.0	−10.0	
14	1064	904	10.5	9.1	19.7	44.1	40.0		4.1		
15	1074	895	10.6	9.0	19.7	42.1	40.0		2.1		
16	1085	886	10.7	9.0	19.7	40.1	40.0		0.1		
17	1085	886	0.0	0.0	0.0	40.1	40.0		0.1		
18	1085	886	0.0	0.0	0.0	38.1	40.0	−2.0	−1.9		
19	1074	895	−10.8	−8.9	−19.7	36.1	40.0	−4.0	−3.9		
20	1063	904	−10.7	−9.0	−19.7	32.1	40.0	−6.0	−7.9		
21	1053	913	−10.6	−9.0	−19.7	26.0	40.0	−8.0	−14.0		
22	1042	922	−10.5	−9.1	−19.7	28.0	40.0		−12.0		
23	1032	932	−10.4	−9.2	−19.6	30.0	40.0		−10.0		
24	1021	941	−10.3	−9.3	−19.6	31.9	40.0		−8.1		
25	1011	950	−10.2	−9.4	−19.6	33.9	40.0		−6.1		
26	1001	960	−10.1	−9.5	−19.6	35.9	40.0		−4.1		
27	991	969	−10.0	−9.6	−19.6	37.8	40.0		−2.2		
28	981	979	−9.9	−9.7	−19.6	39.8	40.0		−0.2		
29	981	979	0.0	0.0	0.0	39.8	40.0		−0.2		
30	981	979	0.0	0.0	0.0	39.8	40.0		−0.2		

Temperature Dynamics

Populations Dynamics
White daisies mass
Black daisies mass

Fig. 6.23 Second Daisyworld simulation

6.11   Summary

We have learned that systems are everywhere, inside and outside us, in the biological, social, and physical world; we started an ideal journey among the systems that characterize:

(a) a typical day in our lives (Sect. 6.1);

(b) our domestic environment (Sect. 6.2), discovering that various forms of direct intervention can be brought to bear on such systems (Fig. 6.1);

(c) the local external *micro* environment (Sect. 6.3), discovering the *overhead control systems* that pervade every moment of our town life (Fig. 6.3);

(d) the external *macro* environment (Sect. 6.4), realizing that even many natural phenomena are governed by control systems in which only natural factors intervene (Fig. 6.8);

(e) the planetary environment (Sect. 6.5), referencing the model of planetary self-regulation in the Daisyworld model (Fig. 6.11); in particular the *Rings* that regulate the Gulf Stream (Fig. 6.10);

(f) the global warming problem (Sect. 6.6 and Figs. 6.13–6.16), thereby adapting the Daisyworld model to the reality of our planet (Fig. 6.12);

(g) our body and articulations (Sect. 6.7) and the postural control systems (Fig. 6.17);

(h) our psychophysical system (Sect. 6.8), in order to demonstrate that we exercise control over our lives in order to satisfy needs (Fig. 6.18) and aspirations (Fig. 6.19);

(i) even the *biological clocks* that maintain the rhythm of our lives are, in fact, natural control systems (Sect. 6.9 and Fig. 6.20).

Chapter 7
The Magic Ring in Action:
Life Environments

The subject of this volume [The Social System] is the exposition and illustration of a conceptual scheme for the analysis of social systems in terms of the action frame of reference [...]

The fundamental starting point is the concept of social systems of action. The interaction of individual actors, that is, takes place under such conditions that it is possible to treat such a process of interaction as a system in the scientific sense and subject it to the same order of theoretical analysis which has been successfully applied to other types of systems in other sciences. (Talcott Parsons 1951, p. 3)

9.1 An organism or society is said to be in dynamic equilibrium if the Self-Assertive and Integrative tendencies of its holons counterbalance each other.
9.2 The term 'equilibrium' in a hierarchic system does not refer to relations between parts on the same level, but to the relation between part and whole (the whole being represented by the agency which controls the part from the next higher level). (Arthur Koestler 1967, p. 347)

In Chap. 6 we headed off on an ideal journey to learn to recognize and model the action of multiple *Rings* operating in various environments. I hope the reader has become aware that without the action of the *Rings*, without their ubiquitous *magic* presence everywhere in a world made up of interconnected variables, life, order and progress could not exist. The aim of this chapter is to guide the reader as he continues this journey through the world of animal or human populations that interact to self-control their dynamics, and as he subsequently moves on to the action of the *Rings* in our ecosystems. I shall then consider the action of control systems in social systems, understood as autopoietic units, which allow such systems to survive, develop their autopoiesis, and maintain in homeostatic equilibrium the relations among individuals and among the basic variables required for their survival. I shall examine in particular the logic of combinatory systems and that of complex adaptive systems (CAS) that support the dynamics within the social systems. The conclusion of this chapter is in line with that of the previous chapters: the biological and social environment cannot survive as a unit, nor can the individuals that compose this environment continue for long to survive, without the action of a countless quantity of interconnected "magic" *Rings*, nested over several levels

and operating continuously. Chapter 8 will then examine organized systems in order to identify the *Rings* that permit the preservation of organizational relations and the interaction among organizations.

7.1   Magic *Rings* in the Macro Biological Environment

Having already examined the *Rings* that permit individuals to exist, let us now examine control systems with regard to *biological environments*, considering above all how the dynamics of any type of biological collectivity can be controlled. I define a *collectivity* as a plurality of *agents* which show an individual micro behavior over time—producing a micro effect—but, considered together, are capable of developing a macro behavior—and/or macro effect—which is attributed to the collectivity as a definable unit.

Collectivities can be composed of persons or animals, active biological organisms or reactive ones (plants, microorganisms); in any case, if considered from a certain distance, from an endogenous perspective (von Bertalanffy, 1968), collectivities can be observed as units (or macro agents) which appear distinct with respect to the individuals they are composed of and show an autonomous observable macro behavior (change in the state of a collectivity over time)—which can lead to observable macro effects—which derives from the interactions of the micro behaviors (change in the agents' states over time) or from the derived micro effects of the agents. Collectivities can be observable (e.g., swarms, flocks, crowds, spectators at a stadium, students in a classroom, people talking in a crowded room, dancers doing the Can Can), or simply imaginable e.g., trailer-trucks traveling a stretch of highway in a month, the noble families of Pavia who erected the 100 towers in the span of two centuries, a group of scientists who dedicate themselves to a branch of research, the consumers of a particular product during its entire life cycle, stockbrokers working on a certain day in world or European stock markets).

Collective phenomena are the essence of life at any level, both for human beings and animals, who operate in populations, societies (cities, nations, tribes, associations, teams, social units) and groupings of various kinds (crowds, hordes, flocks, flights, herds, schools, and so on). I define a *population* as a collectivity of similar individuals (normally of the same species) that live in a given geographical area and can reproduce by crossbreeding with other individuals in the collectivity. Populations can be human, animal, or vegetal. A group of populations living contemporaneously in a given area form a *biological community* constituting an *ecosystem*.

> Since the language of the early ecologists was very close to that of organismic biology, it is not surprising that they compared biological communities to organisms. For example, Frederic Clements, an American plant ecologist and pioneer in the study of succession, viewed plant communities as "superorganisms." This concept sparked a lively debate, which went on for more than a decade until the British plant ecologist A. G. Tansley rejected the

notion of superorganisms and coined the term "ecosystem" to characterize animal and plant communities. The ecosystem concept is defined today as "a community of organisms and their physical environment interacting as an ecological unit". (Capra 1996b, p. 33)

I define s*ociety* (or *social system*) as a collectivity made up of one or more populations, even of different species, that live for a long time (even several generations) in the same territory according to *patterns* of stable *social relations* of coexistence imposed on them (systems of authority created by shared institutions) or which are self-generated (self-organization), thereby permitting coordinated and cooperative behavior for the survival of the society while also allowing its members to reproduce. In many societies the members are *stratified* into homogeneous groups that do not always cooperate. In human societies the members often also share their cultural systems and their expectations for behavior, which is useful and necessary for the general interest (the common good). The above definition is in line with the traditional one offered by sociologists and anthropologists. For example, Sandro Segre writes in his study on the sociologist Talcott Parsons (1951, 1971):

> A "system" is a stable set of interdependent phenomena, provided with analytically-established boundaries, which relates to an ever-changing external environment. A "social system" is a system of social interactions between reciprocally oriented actors. It consists of roles, collectivities, norms and values. The social system forms a system of societies, each characterized by relative autonomy, by its own territorial organization, and by its own sense of identity. In a social system, actors relate to one another by jointly orienting themselves to a situation through a language or other shared symbols. A social system comprises several subsystems or collectivities, all of which are functionally differentiated, interdependent and intertwined. Analysis of such a system focuses on the conditions in which interactions occur within particular collectivities. (Segre 2012, pp. 2–3)

There are other forms of collective behavior that occur in combinatory systems and in CAS. I propose defining a *combinatory system*, or *simplex system*, as any collectivity formed by similar elements or agents (not organized by hierarchical relations or interconnected by a network or tree relationship) which displays an *analogous* micro behavior over time and produces similar micro effects. Together, the micro behaviors and micro effects give rise to a macro behavior (and/or macro effect) attributable to the collectivity as a whole. The collective behavior, though composed of the micro behaviors, conditions or directs the subsequent micro behaviors of the agents following a typical micro–macro feedback process.

Collectivities made up of heterogeneous elements that interact to form dynamic wholes, which are emerging and not predictable *a priori*, shall be defined as *complex systems*. The CAS approach, in particular (Kauffman 1993, 1996; Castelfranchi 1998, Gilbert 1995; Gilbert and Troitzsch 1999) studies how collectivities interact and exchange information with their environment to maintain their internal processes over time through adaptation, self-preservation, evolution, and cognition (in the sense of Maturana and Varela 1980, p. 13) and achieve collective decisions within a relational context of micro behaviors (Rao and Georgeff 1992; Wooldridge and Jennings 1994).

A collectivity of organized agents *specialized* according to function, functionality, functioning, and topology forms an *organization*, a social system (a *social machine* according to Maturana and Varela 1980), where the collective and individual behavior is determined by a network of stable relations, that is, the organization of the agents and processes for the purpose of creating a stable structure.

> *Organization and Structure*. The relations between components that define a composite unity (system) as a composite unity of a particular kind constitute its organization. In this definition of organization the components are viewed only in relation to their participation in the constitution of the unity (whole) that they integrate. This is why nothing is said in it about the properties that the components of a particular unity may have other than those required by the realization of the organization of the unity. (Maturana and Varela 1980, p. xix)

Chapter 8 will deal with the analysis of *organizations* and the *Rings* which allow the former to exist.

7.2 Magic *Rings* That Regulate the Dynamics of Interacting Populations

This *section* presents particular forms for the control of several fundamental variables regarding populations. I shall begin by observing that populations of individuals of a given species—or even an entire species, such as populations that existed in a given period—present at least two interesting trends: a *quantitative* dynamics (i.e., changes in the numbers of individuals over time) and a *qualitative* one (changes in the traits of the individuals over time).

I shall consider above all the *quantitative dynamics* of populations and try to reflect on the fact that the oceans are not only populated by sharks and the savannas by lions, even if these species are among the most efficient hunting machines in their environments; the seas abound in sardines and the savannas are populated by antelopes, even if both of these species are prey for large populations of predators. Above all, populations, and species in general, are normally self-balancing by means of some *natural* control system that stops them from increasing infinitely.

The model in Fig. 7.1 shows the main and general factors that determine the quantitative dynamics of any population. Loop [**R**] in the model in Fig. 7.1 makes clear that every population increases naturally at a net *growth rate*, which is equal to the difference between the birth and death rates. When the birth rate falls below the death rate the population gradually decreases, running the risk of natural extinction. Loop [**B1**] causally links the dynamics of the population to that of the *prey* that provides its food. Under normal conditions, the more the population grows, the more the number of prey and quantity of food decreases, so that the population undergoes an initial balancing. If the prey belongs to another population, there is a joint dynamics between the two populations. If the reduction rate of the prey prevails over that of the natural rate of increase for this population, then the population may even risk extinction. Loop [**B2**] illustrates the same joint dynamics relationship;
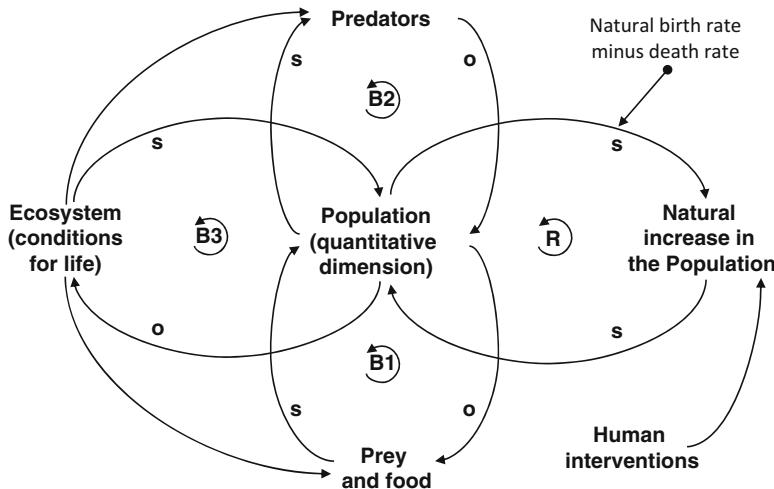
Fig. 7.1 Main factors in population dynamics

however, it considers the observed population as the prey of a population of predators. Taken together, loops [**B1**] and [**B2**] form an initial example of a *food chain*, in which the observed population represents the intermediate *Ring*. Finally, loop [**B3**] shows the relation between the dynamics of the population and the general conditions of life in the *ecosystem*. If the ecosystem offers excellent conditions of life, the population will increase, other circumstances held constant. Normally the opposite trend also holds: the more the population increases, the more the ecosystem deteriorates, as occurs in the ecosystem of the human population. A population that has suffered serious events (fires, floods, etc.) can have fewer possibilities for growth and may also represent easy prey for its predators, in turn encountering difficulties in finding prey or food.

The general model in Fig. 7.1, which can be adapted to every population, even the *vegetable kingdom*, does not illustrate any control system for population dynamics but rather the balancing factors that contrast the reinforcing loop that would lead to natural expansion. Describing and simulating these control systems are difficult tasks, since populations have dynamics that do not appear to approach some limit or "natural" value. Yet even the dynamics of biological collectivities seems to be determined by some control system, even though this is difficult to identify. For many populations man uses the levers available to him to set objectives and limits in order to control their dynamics in an evident manner. Proof of this lies in the hundreds, perhaps thousands, of forms of cultivation and extirpation, of breeding and hunting that exist. In other cases man radically influences the dynamics of populations through forms of protection and tolerance—which leads to overpopulation with respect to the environment (deer, rats)—as well as through actions that threaten their survival (sea turtles deprived of their eggs and whales hunted everywhere, as we shall see in Sect. 7.3).

In other cases populations are regulated by *Rings* that make use of "natural" levers. One such *Ring* includes a very potent lever: the natural instinct of adult males to eliminate a large number of newborns, especially if they are the offspring of other males, as occurs, for example, with hippopotamuses and lions. A similar mechanism can be observed in many other species, though with a variation: often it is the firstborn—for example, in a brood of birds—that stops the other eggs from hatching (Dawkins 1989, p.133). In both cases this control system, though seemingly cruel, ensures that population growth slows down when reaching an upper limit. During some time periods, man has also had to activate some form of self-regulation. In cases of declining population, policies were introduced to spur population growth by activating the levers characterized by the battle against infant mortality, the improvement in living conditions and health, and various forms of economic incentives to produce large families, such as occurred in Italy from the end of 1925 to the end of the Second World War. In other circumstances, where population has increased too much, policies to reduce population growth were undertaken, such as in the People's Republic of China through the introduction of a one-child policy for the control of demographic growth (the laws introduced in 2001 that forbid families from having more than one child) and the tendency of Chinese families to favor male children, especially in rural areas (Von der Pütten 2008). This policy was revised in December 2013, and from 2014 onward Chinese couples will be allowed to have two children.

However, quantitative controls of a population also occur by means of another powerful natural control system represented by conflicts with a second (or more than one) population that are connected to the food chain. When the individuals in a population are the prey of other predator populations, a "natural" control system tends to carry out a reciprocal control of the number of individuals in both populations, often succeeding in producing an equilibrium state or an oscillating dynamics. In order to make it easier to understand the reciprocal control system between the number of individuals in populations (or species) connected in the food chain, I shall only examine two species—which I indicate as **A** (*prey*) and **B** (*predators*)—and consider that, in effect, the evolution of the prey cannot occur without the simultaneous evolution of the predators. There is a coevolution between the two populations. Vito Volterra (1931) and Alfred Lotka's (1925) well-known equations show how two prey and predator populations can have a dynamic equilibrium that takes various forms in relation to the variation rates for the populations, which are a function of the number of individuals in each population (Mella 2012).

The model in Fig. 7.2 describes the simplest equation system, proposed by Vito Volterra (1926), which hypothesizes that prey and predator populations are linearly connected.

Let us assume the observations originate in year $n = 0$, when the amount of prey and predators is determined (we could also have indicated the pairs). Subsequently the populations increase and decrease year after year at a fixed rate. The *prey* (**A**) increase
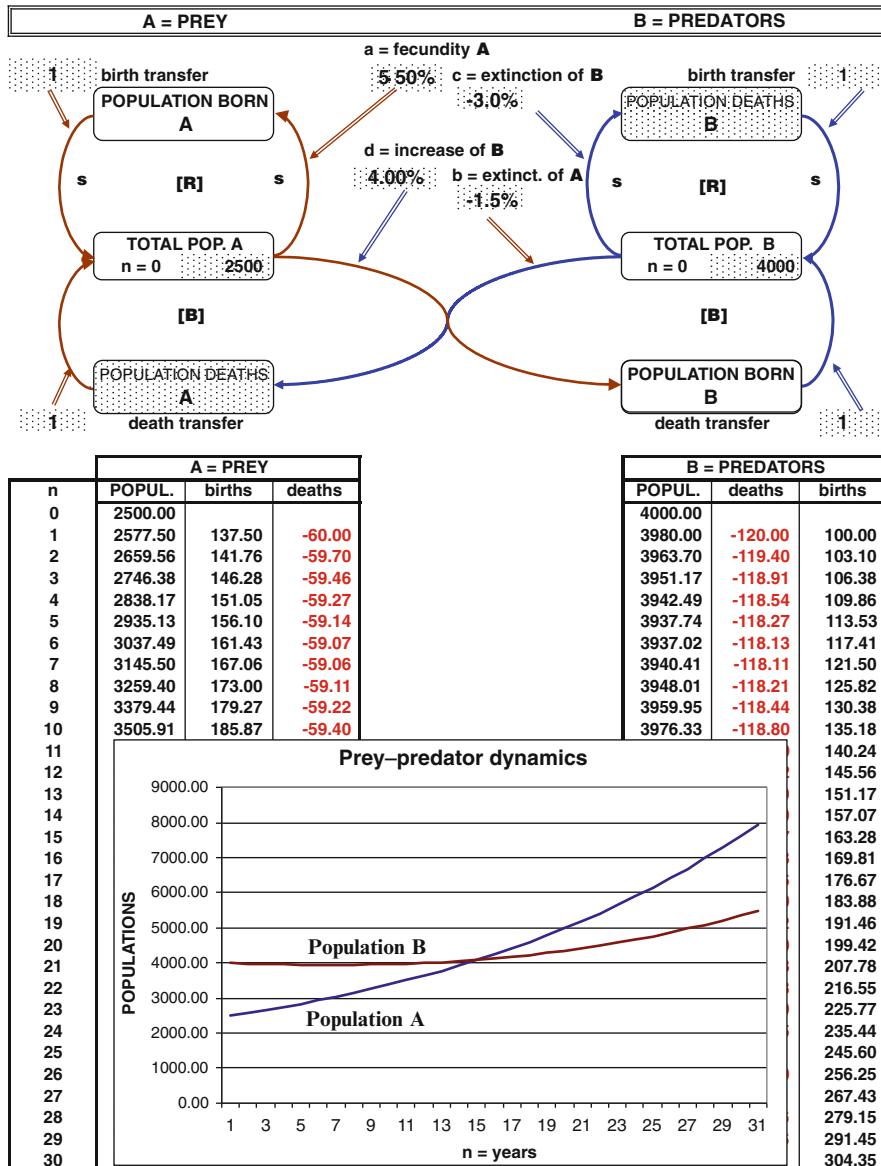
Fig. 7.2 Model of a control system of two populations based on Volterra's equations

by a certain yearly number according to the rate "a" (which takes account of the difference between the natural birth rate and the natural death rate), which is applied to the population of the previous year, A_n, and decrease at a rate "b," which expresses the average voracity of the predators (**B**), taking into account the number of predators from the previous year, B_n. Thus, the prey population at year "$n + 1$" is:

$$\mathbf{A}_{n+1} = \mathbf{A}_n + a\,\mathbf{A}_n - b\mathbf{B}_n \tag{7.1}$$

The *predator* population **B** dies out from natural causes at a death rate "d" and increases by feeding on the prey at a reproduction rate "c," which is tied to the food supply. The predator population in year "$n + 1$" is:

$$\mathbf{B}_{n+1} = \mathbf{B}_n - c\,\mathbf{B}_n + d\mathbf{A}_n \tag{7.2}$$

Together (7.1) and (7.2) form the system that controls the two populations, each a function of the other. The respective dynamics over $N = 30$ periods is shown in the graph in Fig. 7.2. When we insert the data into the simulation model, the model in Fig. 7.2 shows that in year $n = 15$ population **A** = prey reaches and surpasses population **B** = predators, even though it began with fewer individuals.

The reason is clear: there is a positive difference between the birth and death rates, with $(a - b) > 0$, and as a result each year population **A** has a net increase of 4 %; **B**, on the other hand, has a lower rate of increase of $(d - c) = 1$ %.

To make the model more realistic, Volterra proposes a variant—introduced independently by Alfred Lotka as well in his research on the dynamics of the autocatalytic chemical reactions—where the rate of decrease of **A** and of increase of **B** are a function of the number of individuals in *both populations*. In fact, it is intuitively clear that the chance of being the victim of a predator depends not only on the number of predators but also on the number of prey, since a larger population means more likelihood of escaping from the hunt. Moreover, for a predator the possibility of hunting a prey does not depend only on the number of prey but also on the size of the predator population, since the higher the number of predators the more difficult it is to contend for the prey.

The variation in the model consists precisely in setting the number of deaths in population **A** and births in **B** at a critical mass, which is given by the product of the prey and predators: Critical Mass $(\mathbf{A}_n\mathbf{B}_n)$. Thus, (7.1) and (7.2) are modified as follows:

$$\mathbf{A}_{n+1} = \mathbf{A}_n + a\,\mathbf{A}_n - b\,\mathbf{A}_n\mathbf{B}_n \tag{7.3}$$

$$\mathbf{B}_{n+1} = \mathbf{B}_n - c\,\mathbf{B}_n + d\,\mathbf{A}_n\mathbf{B}_n \tag{7.4}$$

These modifications produce the more realistic control system—known as the Volterra–Lotka equations—represented in the model in Fig. 7.3.

As we can see, the "natural" control leads the two populations to maintain a reciprocal relationship that produces an oscillating dynamics, which is, in fact, similar to that of the sardines and sharks in the Adriatic Sea (observed by the zoologist Umberto D'Ancona), which was the basis for Volterra's famous equations.

Other variations have been introduced to "enrich" the system (e.g., the competition for food among predators); since the variants do not modify in any way the

A = PREY   **B = PREDATORS**

	POPULATIONS		PREY		CRITICAL MASS	PREDATORS	
YEAR (n)	PREY	PREDATORS	births	deaths	A B	births	deaths
	A	B	a A	(-) b A B		d A B	(-) c B
0	300.00	100.00			30000.00		
1	301.50	96.70	13.50	12.00	29155.05	5.7000	9.0000
2	303.41	93.54	13.57	11.66	28379.47	5.5395	8.7030
3	305.71	90.51	13.65	11.35	27669.62	5.3921	8.4183
4	308.40	87.62	13.76	11.07	27022.14	5.2572	8.1459
5	311.46	84.87	13.88	10.81	26433.97	5.1342	7.8859
6	314.91	82.25	14.02	10.57	25902.38	5.0225	7.6383
7	318.72	79.77	14.17	10.36	25424.89	4.9215	7.4029
8	322.89	77.42	14.34	10.17	24999.32	4.8307	7.1795
9	327.42	75.21	14.53	10.00	24623.77	4.7499	6.9681
10	332.30	73.12	14.73	9.85	24296.59	4.6785	6.7685
11	337.54	71.15	14.95	9.72	24016.41	4.6164	6.5804
12	343.12	69.31	15.19	9.61	23782.11	4.5631	6.4036

Fig. 7.3 Model of a control system of two populations based on Volterra–Lotka equations

basic logic of the system, those interested in further detail can consult specialist texts (Casti 1985; Takeuchi 1995; Martelli 1999; Flake 2001).

It is difficult to recognize two control systems in Fig. 7.2 and 7.3, since no "objective" (or "limit") is indicated. In fact, these are "natural" control systems—analogous to the one operating in Daisyworld (Sect. 6.5)—where the objective

"reciprocal size of the two populations" depends on the type of population and on the environmental features.

The populations "know" whether or not they can achieve equilibrium and what kind of equilibrium this would be.

> In spite of the extensive knowledge of self-reinforcing feedback in common folk wisdom, it played hardly any role during the first phase of cybernetics. The cyberneticists around Norbert Wiener acknowledged the existence of runaway feedback phenomena but did not study them any further. Instead they concentrated on the self-regulatory, homeostatic processes in living organisms. Indeed, purely self-reinforcing feedback phenomena are rare in nature, as they are usually balanced by negative feedback loops constraining their runaway tendencies. In an ecosystem, for example, every species has the potential of undergoing an exponential population growth, but these tendencies are kept in check by various balancing interactions within the system. Exponential runaways will appear only when the ecosystem is severely disturbed. Then some plants will turn into "weeds," some animals become "pests," and other species will be exterminated, and thus the balance of the whole system will be threatened. (Capra 1996b, p. 63)

The law of dynamic instability (Sect. 1.6.4) is constantly operating. However, in the model in Fig. 7.3 it is possible to exercise an *exogenous control*—for example, by introducing the possibility of varying the numbers in the population through outside intervention—when the size exceeds certain maximum levels or is below given minimum levels. In the case of populations of fish or deer, for example, numbers can be controlled by fishing or hunting, if the number becomes excessive, or by repopulation, if the number becomes too low. Figure 7.4 starts from Fig. 7.3 and completes it by introducing maximum values for the two populations.

The upper limit to the predator population, Pop **B** max = 200 units, has the effect of making the *predator* dynamics regular; however, at the same time it also regulates the dynamics of the *prey* population, which, due to the graduality of the control, reaches a maximum level which initially is slightly above the specific maximum limit, Pop **A** max = 700, subsequently falling to a level below this limit. This is made possible by a *Ring* whose control levers, in this case, consist of a reduction in the populations due to excessive numbers with regard to the objective/limit (the levers are not indicated in Fig. 7.4). In order to make the model even more realistic, we could also assume a modification in the birth and death rates of the populations when the maximum limits for population numbers are approached. It is also useful to consider the natural limits to population expansion from limits to the available resources.

7.3 Magic *Rings* That Regulate Ecosystems

The control system applied to the Volterra–Lotka model for two populations can be extended to any number of populations forming a food ecosystem. In this case we need to define the rules of interaction between populations of prey and predators. As an example, we can consider three populations that form a food chain. The first population of prey, **A**, serves as the first link in the food chain; the second
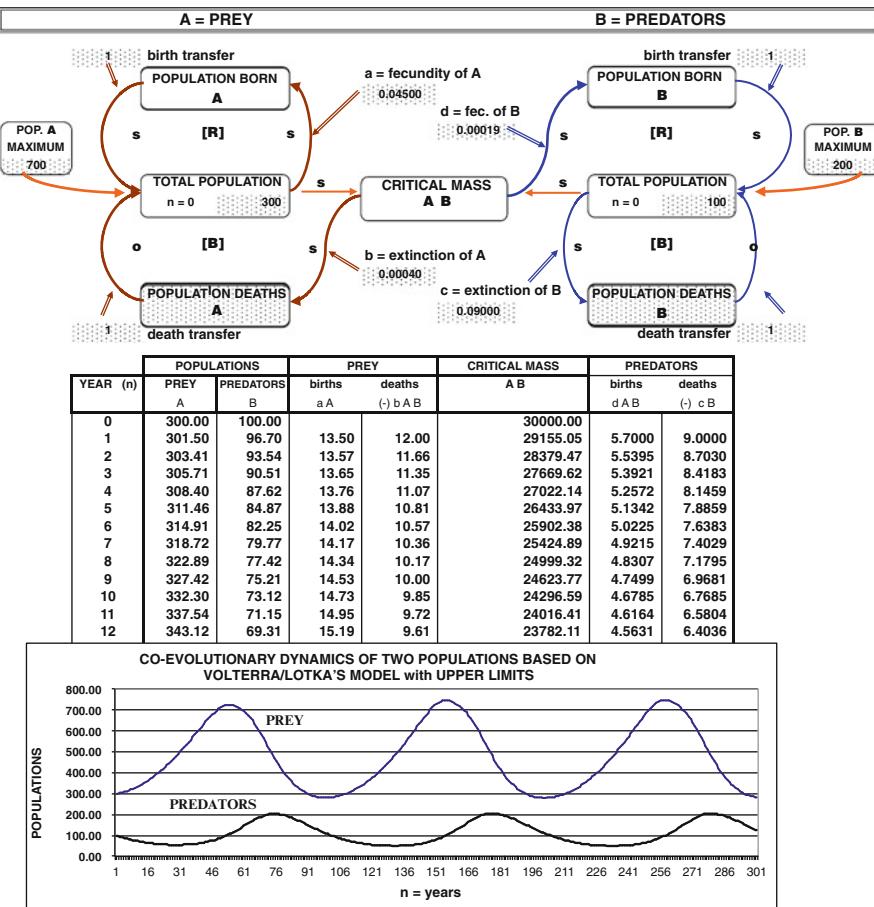
Fig. 7.4 Model of a control system with two populations (Volterra–Lotka), with upper limits

population, **B**, feeds on the prey but represents, in turn, the prey fed on by the individuals in the third population, **C**, which are pure predators that represent the final link in the food chain. Population **C** (pure predators) increases in relation to the available food (population **B**) and becomes extinct because of natural causes, since it is not a food source for other superordinate populations. This situation leads to an annual increase in population **A** at a rate determined by the difference between the birth and natural death rates; moreover, population **A** decreases in relation to the voracity of population **B**. The latter increases and becomes extinct as a function of populations **A** and **C** (prey and predators, respectively) at rates that are net of births and natural deaths.

Figure 7.5 shows the complex oscillatory dynamics that result from the assigned parameters over 1,000 periods—months or years—after having also set *maximum*
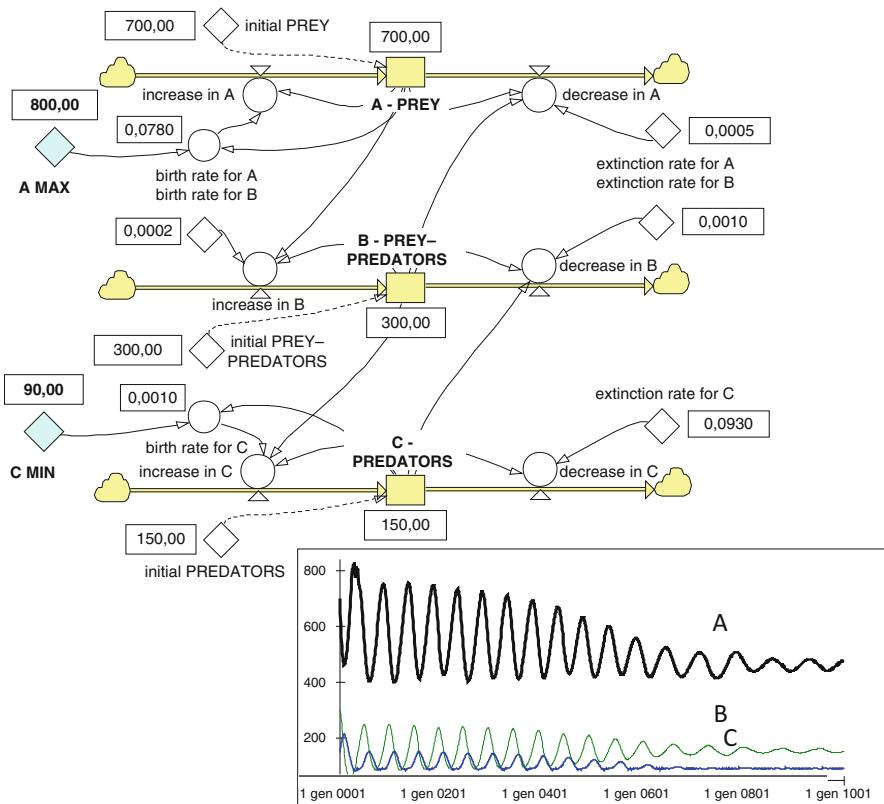
Fig. 7.5 Volterra–Lotka model for three populations, simulated using Powersim

limits for the prey and *minimum* ones for the predators limits that, representing merely an example, influence only the birth rates of the two populations.

From the graph we can see that the dynamics of prey **A** conditions that of population **B**, which presents oscillations that tend toward stable values due to the effect of the control. We could easily set more elaborate limits that call for a maximum amount of resources available or territory size, referring to these to specify the birth and death rates and the outside interventions.

The same logic applied to the simulation of three populations can be used for any number of populations. Figure 7.6 illustrates the model created with Powersim to simulate the dynamics of five populations for 5,000 periods.

Figure 7.7 presents the results of the simulation. The top graph shows the population dynamics determined only by the natural *Rings* of reciprocal control carried out by the populations forming the food chain. The other graph presents the dynamics produced when we impose (for simplicity's sake) maximum constraints on populations **A** and **B**.

The upper graph in Fig. 7.7 shows that the irregular, probably chaotic, dynamics that occur using only the reciprocal natural *Rings* (lower graph) becomes regular as
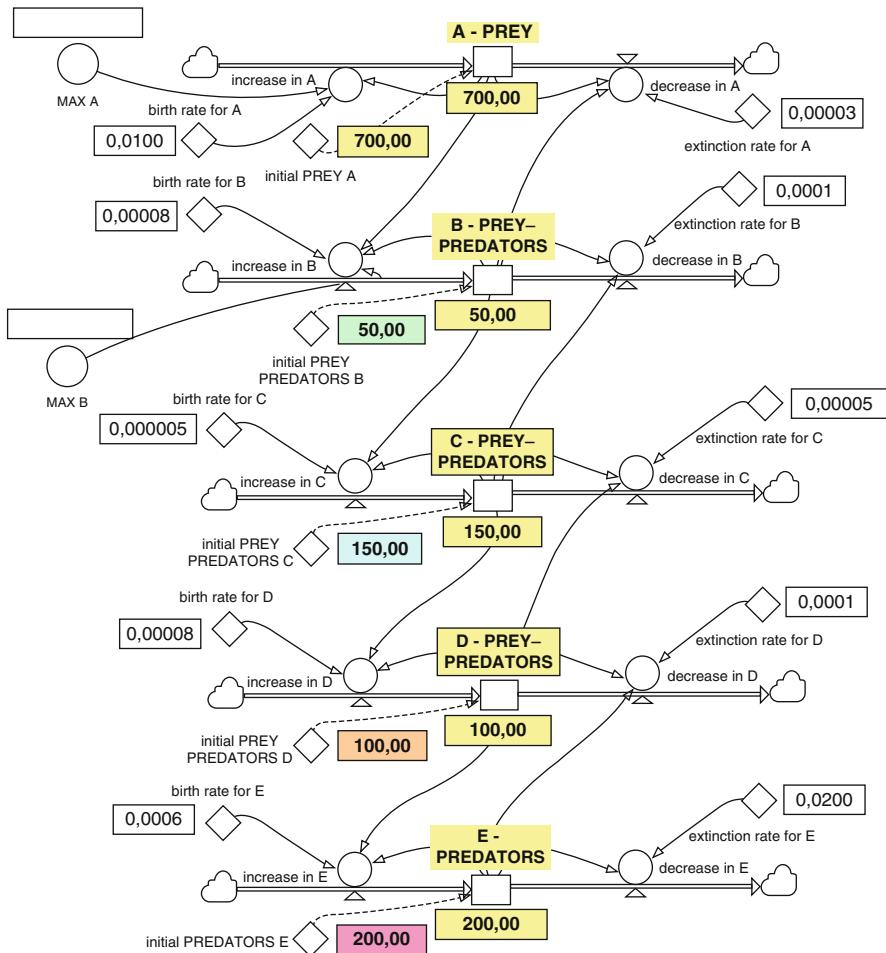
Fig. 7.6   Model of an ecosystem of five populations forming a food chain

soon as the simple external control hypothesized in the simulation is carried out. In particular, in order to eliminate the fluctuations in population dynamics and obtain the more regular dynamics shown in the lower graph, only two limits (not indicated in Fig. 7.6) have been introduced in the simulation for simplicity's sake: Max population $A = 2,000$; Max population $B = 800$.

   We need to be aware that *ecosystems* are very difficult to simulate, since they can include thousands of populations of different species (e.g., a rain forest or coral reef), whose reciprocal relations are so complex that it appears impossible to translate them quantitatively into birth and death rates and predation.

   Ecologists have attempted to overcome the inadequacies of their simple models by including a structured hierarchy of species that are referred to as "food webs." In such a hierarchy there is a pyramid that is surmounted by the top predator, such as a lion, with the smallest
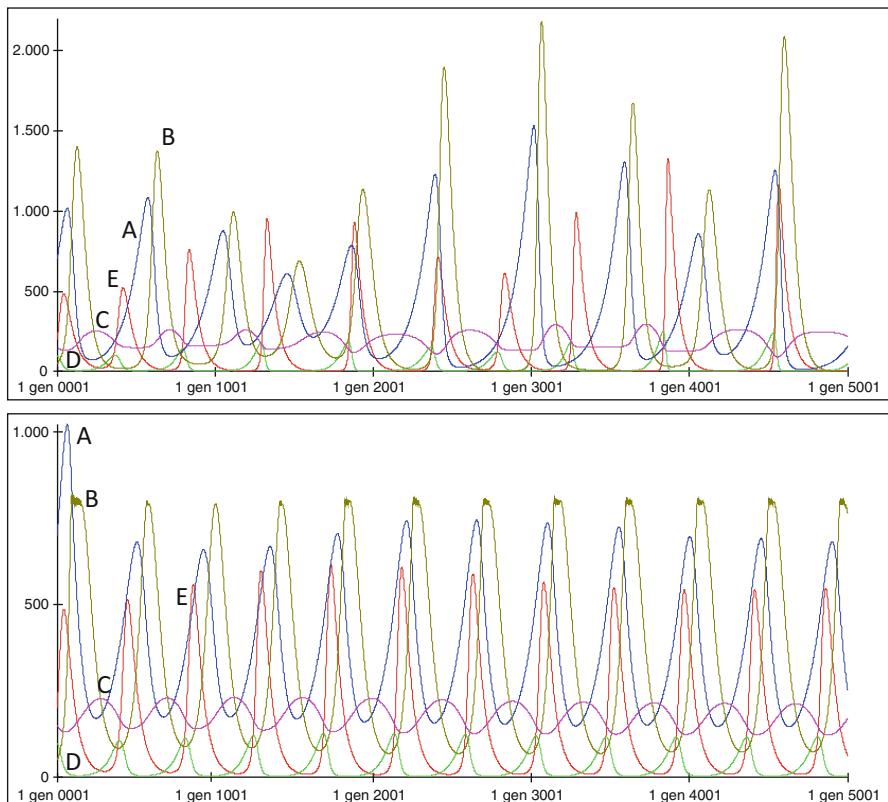
Fig. 7.7 Dynamics of five populations of an ecosystem, without (*top graph*) and with (*lower graph*) external control

numbers. The numbers increase as you go down each "trophic" level, until at the base of the pyramid are the most numerous primary producers, the plants, that provide food for the whole system. In spite of years of effort and computer time, the ecologists have made no real progress towards modeling a complex natural ecosystem such as a tropical rain forest or the three-dimensional ecosystem of the ocean. No models drawn from theoretical ecology can account in mathematical terms for the manifest stability of these vast natural systems. (Lovelock 1988, p.49)

7.4 Magic *Rings* That Regulate the Qualitative Dynamics of Populations Over Time

Let us now also consider the qualitative dynamics of populations—more generally, species—limiting ourselves to the control process carried out by natural selection, a powerful "natural" control system that leads to the gradual modification (in terms of improvement) of the *phenotypic traits* of the individuals in a population that give
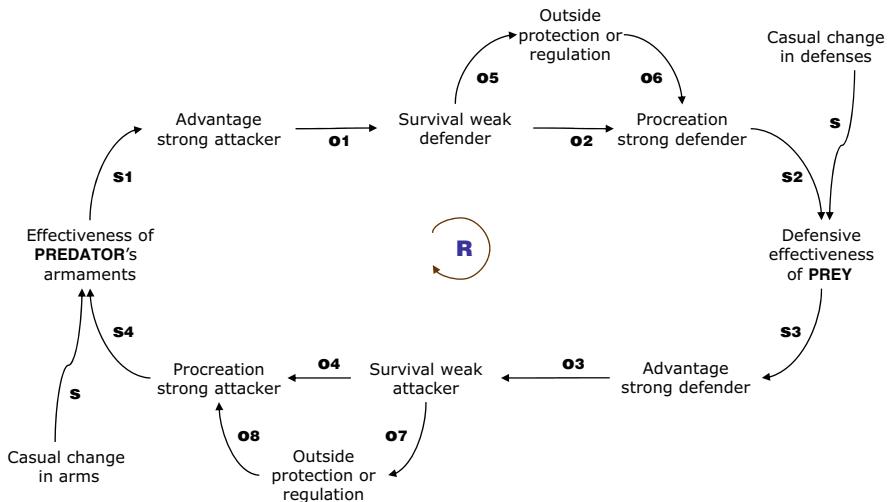
Fig. 7.8 Model of the evolution of phenotypes based on the process involving the struggle for life

them the maximum chance of survival. Not wishing to present too simplified a view of evolution, I shall limit myself to citing the concise thoughts of Darwin about the natural selection process based on the struggle for survival among individuals, which leads to the qualitative evolution of the species through the "natural selection of the fittest"—whatever interpretation we wish to give to this process, including Dawkins' version (1989, 1st edn., 1976) of *selfish gene*—as illustrated in the model in Fig. 7.8. The model shows the dual facts that while the life of a predator (attacker) depends on its offensive weapons for killing its prey, the life of the prey (defender) is linked to the predator's defensive capabilities.

Predators with refined weapons (whatever the interpretation we choose to give to this term) have an advantage in their predation (s_1), thereby reducing the chances for survival for the weaker prey (o_1) (defenders); the lower number of weak prey increases the probability of procreation (o_2) by the prey with strong defenses (hereditary transmission of traits), which leads to an increase (s_2) in the effectiveness of the defenses for an increasingly higher number of prey. The descendants of the stronger prey have a high probability of inheriting the effective defense capabilities of the parents; thus, an ever higher number of prey can escape from their predators (s_3). The strengthened defense capacity of the prey reduces the probability of survival of the weaker predators (o_3) (attackers); with more effective offensive weapons, the remaining predators can reproduce with greater frequency and with a good chance of passing on their weapons to their offspring (o_4). On average the predators have become stronger (s_4) and can thus eliminate the weaker prey, thereby ensuring that those with more effective defense capabilities can procreate and spread these capabilities to their descendants. The loop repeats itself, generation after generation (the system is repetitive), thereby producing a gradual improvement in the predators' weapons (a lengthening beak, more powerful jaws,

greater speed of attack, etc.) and in the defense mechanisms of the prey (improved mimesis, thicker shells, greater speed of escape, etc.).

This process is in line with the masterful description by Darwin himself, which it is useful to reproduce in full:

> Owing to this struggle for life, any variation, however slight and from whatever cause proceeding, if it be in any degree profitable to an individual of any species, in its infinitely complex relations to other organic beings and to external nature, will tend to the preservation of that individual, and will generally be inherited by its offspring. The offspring, also, will thus have a better chance of surviving, for, of the many individuals of any species which are periodically born, but a small number can survive. I have called this principle, by which each slight variation, if useful, is preserved, by the term of Natural Selection, in order to mark its relation to man's power of selection (Darwin 1859, p. 61).

> Natural selection will modify the structure of the young in relation to the parent, and of the parent in relation to the young. In social animals it will adapt the structure of each individual for the benefit of the community; if each in consequence profits by the selected change. What natural selection cannot do, is to modify the structure of one species, without giving it any advantage, for the good of another species; and though statements to this effect may be found in works of natural history, I cannot find one case which will bear investigation. [...] Sexual selection is, therefore, less rigorous than natural selection. Generally, the most vigorous males, those which are best fitted for their places in nature, will leave the most progeny. (Darwin 1859, pp. 86–87)

I would also suggest everyone meditate on the vision of evolution proposed by Dawkins (1989, 2004), which places at the center of evolution the search for stability in the *genes* that transform the phenotypes of organisms (their *survival machines*) in order to ensure greater "*chances of survival*" for their offspring in subsequent generations.

> Natural selection in its most general form means the differential survival of entities. Some entities live and others die but, in order for this selective death to have any impact on the world, an additional condition must be met. Each entity must exist in the form of lots of copies, and at least some of the entities must be potentially capable of surviving – in the form of copies – for a significant period of evolutionary time. Small genetic units have these properties: individuals, groups, and species do not (Dawkins 1989, p. 33).

> The gene is defined as a piece of chromosome which is sufficiently short for it to last, potentially for long enough for it to function as a significant unit of natural selection. Exactly how long is 'long enough'? [...] It is its potential immortality that makes a gene a good candidate as the basic unit of natural selection [...] Genes are competing directly with their alleles for survival, since their alleles in the gene pool are rivals for their slot on the chromosomes of future generations. Any gene that behaves in such a way as to increase its own survival chances in the gene pool at the expense of its alleles will, by definition, tautologously, tend to survive. The gene is the basic unit of selfishness (*ibidem*, pp. 35–36).

> We are survival machines – robot vehicles blindly programmed to preserve the selfish molecules known as genes. This is a truth which still fills me with astonishment. (*ibidem*, p. xxi)

Whatever idea we have of the evolutionary process, the *quantitative* dynamics of populations and species also influences the *qualitative* dynamics; an alteration in some traits of a species can produce considerable changes in the number of individuals in that and other species connected to it in the ecosystem, at times with catastrophic effects (Buchanan 2000).
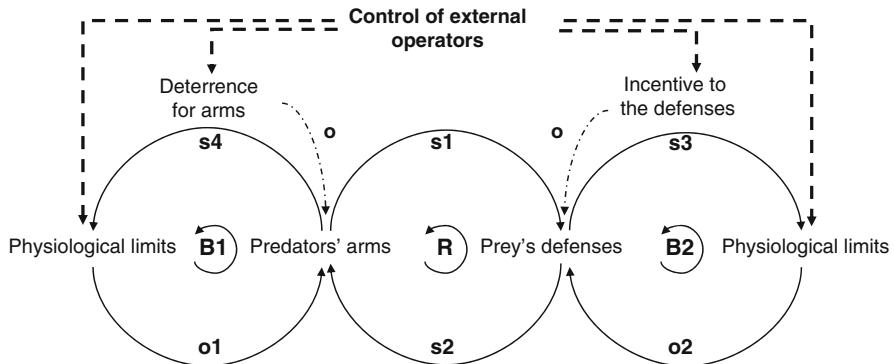
Fig. 7.9 Natural and external Rings regulating the evolution of phenotypes

The dynamics shown by the model in Fig. 7.8 provide insight into these processes; I have added two loops to loop [**R**] in the model—"o_5" and "o_6," as well as "o_7" and "o_8." These represent two forms of *external control* of the process which *weaken* the natural selection activity since, by "short-circuiting" the two relations that are delimited by these loops, they reverse the gradual expansion of the offensive and defensive weapons, thereby making both the prey and predators increasingly weaker and "undefended" in their struggle for survival. I will not comment further, only to say that this weakening mechanism acts whenever human intervention tends to protect some species (first of all the human species) in the struggle for survival, thereby causing individuals to no longer be suited to competition; in fact, we can note the difficulties animals that have been held in captivity (protected) for a long time experience when they are reinserted into their natural habitat. Nevertheless, we must observe that the loops "o_5" and "o_6," as well as "o_7" and "o_8," can also act in the opposite way: when there is an absence of external protection the winning traits are reinforced. This is the principle all breeders follow when they "guide" the reproduction of those individuals with the strongest traits.

Let us return to the main loop, [**R**], and ask whether, if left alone, evolution leads to a strengthening without limits in the attack and defense mechanisms—which also holds for the traits that provide an advantage in reproduction—or if, once again, some "natural" control system intervenes to limit the excess numbers. The reader should go back to the model in Fig. 1.12, which represents the general law of *dynamic instability*: nothing grows or remains stable endlessly. Stability is always accompanied by instability. This general law demonstrates that all growth—just as all declines—will be halted at a limit determined by the operation of some control system. The "natural" control system that sets a limit to the growth in attack and defense organs in the world of biology is represented—in simplified form—in Fig. 7.9.

When the weapons of the predator increase, there is soon a slowdown in the chances for further growth due to the presence of physiological limits (o_1), in accordance with loop [**B1**]; when the limit is reached, the individuals are incapable

of surviving and therefore die, without offspring, thereby stopping any further transmission of their traits (s_4). The same process regulates the development of defenses in the prey, as shown in loop [**B2**].

I observe that if we substitute [predator weapons and prey defenses] with [nation A's weapons and nation B's weapons] (where A and B are in a perennial state of mutual tension), [power required by new software and power of new computers], [development of new antibiotics and resistance to antibiotics], [equipment of police forces and the slyness of criminals], or [computer virus defenses and the insidiousness of new viruses], we will always obtain the same logical structure of this control system.

One final observation: the model in Fig.7.8 would invariably operate even if the conditions of the ecosystem in which the population is living should change (Fig. 7.1). Let us assume that a sufficient number of predators and prey (e.g., lions and gazelles) find themselves, due to a cataclysm, living in a different environment (e.g., a mountain area without level ground). After a sufficient number of years we would probably find gazelles and lions capable of climbing and jumping. Loop [**R**] would invariably operate to heighten these capacities in both populations; in the absence of mutual adaptation, one or both populations would become extinct.

7.5   The *Rings* Regulating Social Systems: The Control of Coexistence

Even if the distinction between the *social*, *economic*, and *political* aspects of our behavior is relative, the *social environment* of which we all are a part is made up of a complex system of groups of individuals, of expanding size—individuals, couples, families, tribes, clubs, clans, villages, regions, populations, nations, associations, and organizations of various types and sizes—that mutually condition one another through a network of social and economic connections surrounding consumption and deaths (destruction), production and births (regeneration), and the exchange of persons, resources, labor, information, etc., which permits the replacement of elements (persons, resources, labor, information) that are no longer efficient with new elements, thereby indefinitely maintaining the network of connections.

Because there are no other factors that determine the maintenance of the social system apart from the network of connections which regenerates the elements forming those connections, this is a clear example of *autopoiesis* operating at a social level, similar to what Maturana and Varela indicated in their definition of a living system. In fact, Maturana and Varela stated that a social system can be defined as a structured system of autopoietic systems made up of individuals linked by social and economic relationships who seek to maintain their autopoiesis through social interaction.

> The question, 'What is a social system?' cannot be answered by simply describing a particular one because we do not know the significant relations that we must abstract when characterizing its organization. The question must be answered by proposing a system which, if allowed to operate, would generate a phenomenal domain indistinguishable from the phenomenal domain proper to a natural social system. Accordingly, I propose that a collection of autopoietic systems that, through the realization of their autopoiesis, interact with each other constituting and integrating a system that operates as the (or as a) medium in which they realize their autopoiesis, is indistinguishable from a natural social system [. . .]The structure of a society as a particular social system is determined both by the structure of its autopoietic components and by the actual relations that hold between them while they integrate it. (Maturana and Varela 1980, pp. xxiv–xxv)

The authors also specify that if a social system can also include different autopoietic systems, each of these can be part of various social systems.

> An autopoietic system participates in the constitution of a social system only to the extent that it participates in it, that is, only as it realizes the relations proper to a component of the social system. Accordingly, in principle, an autopoietic system may enter or leave a social system at any moment by just satisfying or not satisfying the proper relations, and may participate simultaneously or in succession in many different ones. (Maturana and Varela 1980, p. xxv)

When the *social system* is made up of multiple autopoietic systems, populations or social groups, these can *cooperate* in order to achieve the *common goal* of maintaining the collective autopoiesis, that is, the *common good* of the components of the *social system*. On the contrary, they can also be *in conflict* with other social systems perceived as a threat or obstacle to the achievement of individual autopoiesis.

If the socioeconomic systems are as a whole *autopoietic* and made up of a multitude of component autopoietic systems that tend toward the *common good*, then it is clear the social systems are *homeostatic machines* "*. . .whose own organization (network of defining relations) is the fundamental variable that it maintains constant.*" In other words, the systems must regulate or maintain constant the *network of relations* that defines its own organization.

> Thus small, closely knit communities have a very considerable measure of homeostasis; and this, whether they are highly literate communities in a civilized country or villages of primitive savages. Strange and even repugnant as the customs of many barbarians may seem to us, they generally have a very definite homeostatic value, which is part of the function of anthropologists to interpret. (Wiener 1961, p. 160)

It is thus intuitive that, being *homeostatic* systems, *social systems* can produce their autopoiesis for a long time thanks to the action of the innumerable *Rings* which, at any moment, allow social interactions to develop, become coordinated, and constantly reproduce over time. It has long been recognized that control systems play an indispensable role in maintaining social systems. In his studies on social systems, the sociologist Kingsley Davis clearly recognizes that no social system can survive without a system of norms to which the individuals must conform:

> Always in human society there is what may be called a double reality – on the one hand a
> normative system embodying what ought to be, and on the other a factual order embodying
> what is... These two orders cannot be completely identical, nor can they be completely
> disparate. (Davis 1949, p. 52)

The norms that lead to this factual order act as standards of behavior, or as true
objectives of the social controls:

> [Norms] are controls; it is through them that human society regulates the behaviour of its
> members in such ways that they perform activities fulfilling societal needs – even some-
> times, at the expense of organic needs. (Davis 1949, p. 52)

The logical consequence of the above aspects is that social systems must be
characterized by *three structural elements*.

1. A stable *system* of *social and economic rules* of coexistence (objectives,
   constraints, limits, even in the form of obligations and prohibitions) must
   circumscribe the behavior of the individuals and groups that make up the social
   system (in order to allow them to achieve their objectives without harming or
   destroying the homeostatic relationships that guarantee survival), so that the
   autopoietic relations are not altered, or, if altered, are restored, thereby allowing
   the groups and organizations to function as an autopoietic homeostatic machine.
   A social system made up of systems of individuals that share the same system of
   rules, including language, religion, traditions, and governance, takes the form of
   a *nation*.
2. A stable *network* of *Rings* must make such a system of rules cogent, considering
   the rules as objectives, constraints, or limits which human behavior must
   achieve/respect and maintain. Control systems must identify the in behavior
   that goes beyond the rules and activate the most appropriate levers to bring
   behavior back in line with the rules.
3. A stable *structure* of *social institutions* must guarantee the functioning of the
   stable network of control systems, acting as the governance that establishes the
   rules for coexistence and as the management that regulates the control levers.
   These institutions must guarantee the existence of efficient *chains of control*,
   that is, of effectors, detectors, and regulators that permit the concrete achieve-
   ment of the *Rings* needed for the coordinated regulation of the rules which are
   "vital" for the long-lasting achievement of homeostasis. I propose defining these
   three elements—rules, control systems, and the institutions that produce them—
   as a *political system* which complements and completes the *social system*. The
   structure and the *modus operandi* of the *political system* depend on the cultural
   systems of the individuals or groups that make up the social system.

Taking account of the inevitable and fundamental economic interactions, I shall
refer to the social system by the more complete term *socioeconomic political
system*. When legalized in a stable manner, the *socioeconomic political system*
referring to a given territory becomes a *state*, and the individuals living there, and
who are legally subject to its norms, can be referred to succinctly as the *people* of
the state. The logic of control systems which maintain individual behavior in line
with the rules of coexistence can be outlined in the model in Fig. 7.10.
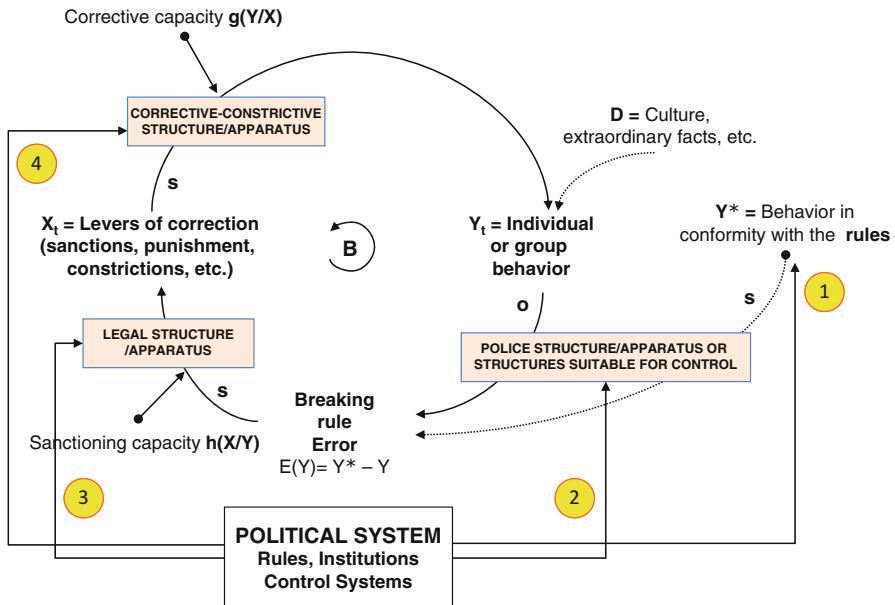
Fig. 7.10 The *Ring* for the control of the respect for the rules of coexistence in the social system

In this figure the *rules*, ①, which discipline the individual behaviors of coexistence are usually translated by the *political system* into the formal laws of the legal system of the social system, or are represented by the "natural" laws codified as normal behavior or that derive from custom. The two types of rules interact in a network that is often difficult to identify; nevertheless, to make the action of control systems concretely possible, the institutions tend, where possible, to transform the "natural" rules into formal legal norms. It seems incredible when we consider how many laws, limits, bans, prescriptions, or obligations regulate our behavior at every moment of the day. For example, our home and the use of public services; traffic laws, regional and municipal regulations, as well as tax laws discipline our every movement in our towns, from the use of our car to that of public transport, from driving to parking regulations, automobile taxes to tolls. Norms, regulations and customs of all types regulate how we enjoy our free time, imposing obligations and prohibitions in all areas: from the bar to the canteen, the church to the beach, weddings to funerals. And norms regarding ceremonies and proper etiquette are no less stringent.

As shown in Fig. 7.10, in order to allow the social *Rings* to operate, the institutions must also create the following necessary apparatuses of the *chain of control*:

– *Detectors*, ②, to identify and recognize behavior not in conformity with the rules: police apparatuses or other structures to check on the conformity of behavior in various areas: judges, arbitrators, head masters in schools, auditors, etc.;
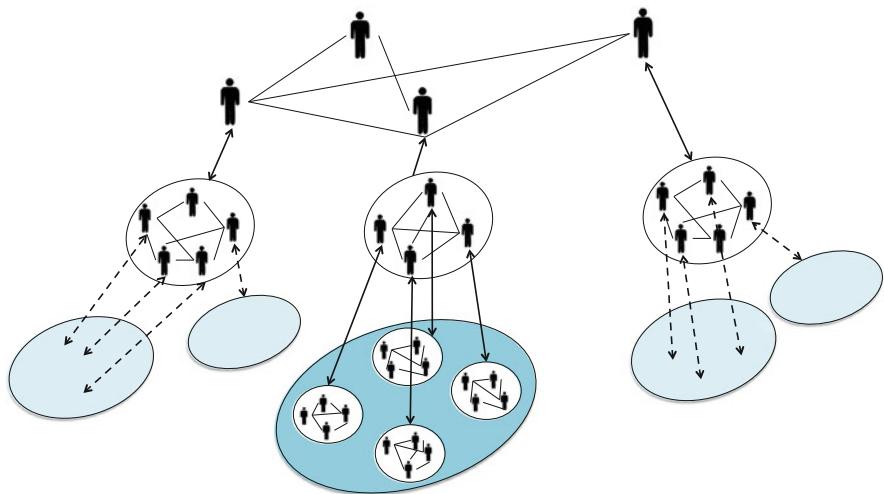
Fig. 7.11 Institutions over various layers, forming a holarchy

- *Regulators*, ③, of the behavioral control levers (the legal apparatus and other apparatuses capable of deciding on types of sanctions);
- *Effectors*, ④, which bring nonconforming behavior back in line or avoid future nonconforming behavior (corrective and, if necessary, constrictive apparatuses, even entailing limits on personal freedom).

The institutions that define the rules and produce the network of control systems are the *governance institutions*, which, within a *state*, represent the *government* structures. The *political institutions* that make up the *government* are normally holonic in nature, since they are arranged in various layers that form a more or less complex holarchy, in order to identify the *hierarchy of control systems* and the *managers* that will guarantee they function properly, as shown in the model in Fig. 7.11.

Such hierarchical forms of *government* arise from the need to control individuals or groups belonging to an increasingly larger territory. The holonic institutions can be either *ascending* or *descending*; if the base individuals (base holons) that form the "people" belonging to the social system produce first-level institutions which, in turn, form second-level ones, and so on, up to the maximum level (top holon), then the process is *ascending*. If the highest level (top holon) is formed, to which maximum authority is attributed, and this level then produces, through a division of power, the lower levels, the process is *descending*, as illustrated in the diagram in Fig. 7.12.

In general, we can state that the number of levels of institutions is higher the vaster is the social system (large territory) and the more need is felt for a control of individuals' behavior (conflictuality); nevertheless, we can also observe a tendency for a reduction in the levels of the governance system and an increase in the number of individuals and in the size of the territory governed by the elements at
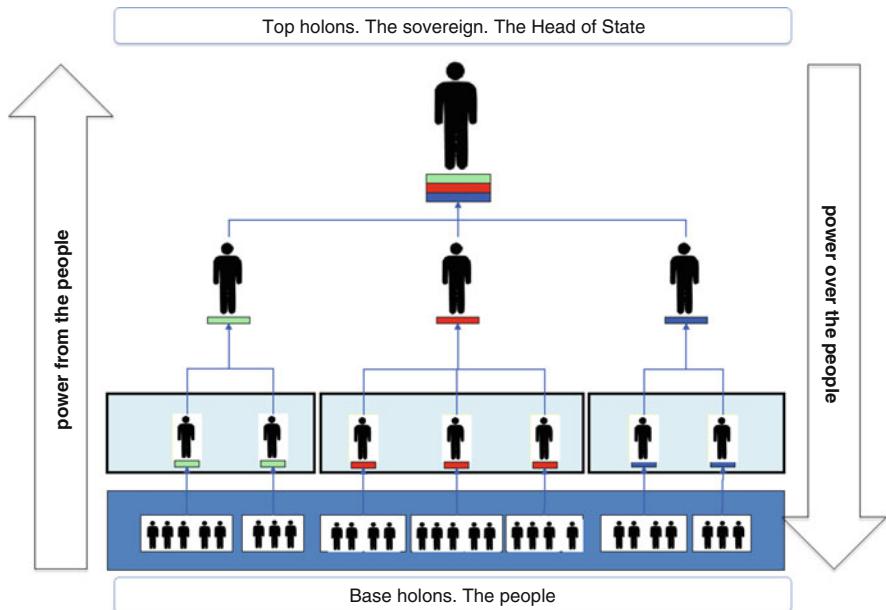
Fig. 7.12 Ascending and descending institutions. The different dimensions in the various levels indicate the concentration of power

each level. A social system comprised of a tribe has, as its highest institution, the tribal chief, who controls the behavior of the tribe's members in order to avoid homicides, thefts, infidelity, as well as to impose various types of tributes. The tribal head can be elected (ascending) or can dominate the "people" through physical or religious force (descending).

When the social system grows through the union of different tribes, a sovereign is needed who can control the tribal heads that, in some manner, must be maintained. When the system grows further, an emperor is required who can control the various sovereigns, however they are referred to. The formation of high-level governance structures (empires or the political union of states) is a result of the need for the control and governance of increasingly larger territories (the Roman Empire, the Chinese Empire, the Mogul Empire, the Ottoman Empire, the USA, USSR, EU, etc.). Many of these institutions are ascending in nature, elective (the tribes elect the tribal heads, who then elect the sovereign, etc.) but history teaches us that institutions that develop through an ascending process very often are transformed into descending institutions. The emperor is not elected but acquires authority and power from some "supreme authority," and in turn appoints the hierarchy of subordinate controllers, who have authority over individuals of increasingly smaller territories. Consider, for instance, the levels of the aristocratic hierarchy in Europe: emperor, king, prince, grand duke, archduke, landgrave, duke, marquis, count, viscount, knight, and noble.

Today the descending institutions of monarchies and dictatorial regimes are still with us; however, more numerous are the ascendant republics or parliamentary monarchies, which arise and are maintained in social systems where the "people" hold power, which they delegate in various forms and for more or less lengthy periods, to their delegates. In many countries today there is a hierarchy of delegated authority in cities, provinces, regions, nations and supranational organizations. This delegated authority results in a *governance* that establishes the policies based on some rule of representation of the delegates; these rules can at times make the governance unstable and the setting of objectives and policies problematic. We can easily find numerous examples of this in the daily news.

It is not beyond reason that, with progress in systems for the exchange of information through networks (the web), we may also someday witness a "flat" governance where the intermediate levels are eliminated, as some have already hypothesized.

> This allows us to conceptually divide the system into functional components responsible for the different stages of the processing of incoming matter and energy (metabolism), and for the processing of information needed to maintain cybernetic control over this mechanism (nervous system). As the system continues to evolve, on-going adaptation and division of labor lead to an increasingly diverse, complex, and efficient organization, consisting of ever more specialized components. This general model of complex, self-organizing systems can be directly applied to the present development of society. Since society is an organismic system consisting of organisms (individual people), it can be viewed as a "superorganism". Conspicuous trends such as globalization, automation, and the rise of computer networks can be understood as aspects of the general evolution towards increasing efficiency and interconnectedness which makes the superorganism ever more robust. (Heylighen 2013, p. 32)

7.6 *Rings* That Maintain Autopoiesis in Social Systems

The control of the coexistence among individuals and groups—that is, the conformity of behavior to social rules—is necessary, though not sufficient, for the integrity of the "social fabric." Equally essential is the maintenance of *autopoiesis* in the social system as a whole. Thus, we must join other *macro* systems of *overall* control to the *microRings* that regulate individual relationships, guaranteeing in this way the regulation of the fundamental socioeconomic *macro variables* which characterize the *common good*. It is up to the *political system* to define the *objectives* of these autopoietic macro variables and determine the order in which the objectives must be pursued, thus performing the fundamental function of *governance* required to activate the macro *Rings* which, by controlling those variables, guarantee autopoiesis in the social system.

In order to control these socioeconomic *macro variables*, government institutions normally create appropriate *institutional macro structures* that are functionally specialized in managing *control systems* in order to achieve the *macro objectives* (the policy) defined by the governance. The totality of these structures represents the *operational apparatus* of the social system, on whose efficiency and
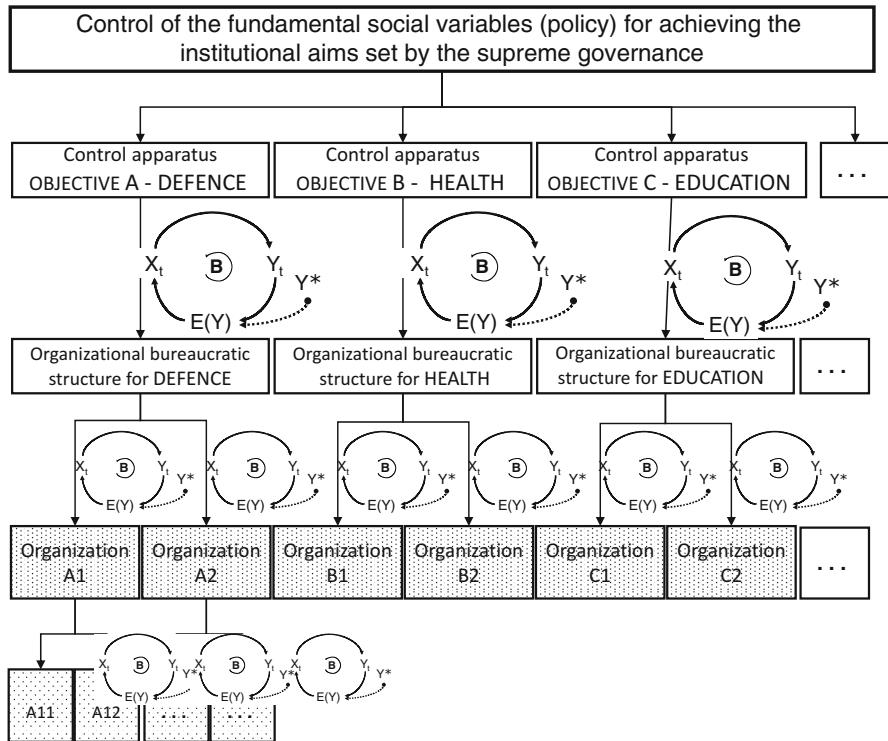
Fig. 7.13 Structure of the organizational bureaucratic apparatus that manages the *Rings* that control the autopoietic variables based on the governance policy

size the operational efficiency of the control systems depends. The specialized organizations that permit the actual functioning of the *operational apparatus* make up the *bureaucratic-administrative apparatus*.

The *macro structures* are at the top of a holarchy because they are operationally divided into smaller structures with more circumscribed control powers, arranged in layers of *bureaucratic-administrative micro structures*, so that even the macro objectives can be holonically detailed as lesser objectives that form a holarchy of objectives, for whose achievement holarchies of control systems are formed.

In a modern social system, with ascending governance, the *bureaucratic-operational holarchies* are usually specialized for each of the social macro variables whose macro-objectives are set by the governance (policy), as shown in the model in Fig. 7.13.

In every modern state, the *bureaucratic-administrative* structures at the highest level are the ministries (however denominated), each of which represents the management for the control of the social variable assigned to them. The objectives the ministries must achieve are divided into segments of objectives assigned, for example, to the ministry heads, who in turn control the increasingly smaller territorial units. In this way, offices, agencies, committees, etc., are created that

SOCIAL VARIABLES [SAMPLE LIST]	GENERAL OBJECTIVES ASSIGNED TO THE SOCIAL VARIABLES [SAMPLE DESCRIPTION]
Defence	Maintain security in the face of external threats or aggression to the state
Public order and internal security	Guarantee adequate levels of internal security by controlling public order and dealing with emergencies and calamities of all kinds
Justice	Provide maximum levels of criminal, civil and juvenile justice services along with adequate levels of prison facilities
Public health	Assure the highest levels of heath care and of treatment and prevention of illnesses
Education (schools, universities)	Assure the highest levels of educational services at all levels, from pre-school to doctoral
Economy	Provide incentives to, coordinate and develop all forms of economic activity
Industry, commerce, agriculture and tourism	Regulate and develop the main forms of production activity
Employment	Regulate the demand and supply of labor by increasing employment
Finance and treasury	Contain and maintain in equilibrium the flows and levels of tax revenue and expenditures (taxes, public debt and public expenditures)
Families	Assure the creation and maintenance of families of all forms
Sport	Assure adequate access and support to all forms of sport
Protection of territories and seas	Further environmental protection, economic growth and social cohesion in the territories
Infrastructure and transport	Provide adequate and modern systems of infrastructure and of public and private transport. Etc.

Fig. 7.14 Social variables necessary for the maintenance of autopoiesis

provide the local governance and serve as elements in the chain of control of the various *Rings*.

In every society we encounter the same higher-level "autopoietic variables" to which the governance *policies* assign objectives and priorities. In order to identify these variables, it is sufficient to make a list of the ministries that, though under different names, every state must set up. Limiting ourselves to only the most general and obvious ones, the fundamental variables that every state tends to control, and whose functions are assigned to the ministries, can be indicated in the table in Fig. 7.14.

Thus, for example, the macro variable "*public health*" is assigned general objectives by the general governance; these are usually qualitative objectives (health care for everyone, increase in the number of hospitals, a reduction in death rates by y %, etc.), which are then divided into territorial objectives assigned

to the second-level governance, and so on for each successive territorial unit down to the cities and, even more specifically, each health center. For each of these objectives, each territorial level, there are *Rings* operating to achieve them, whose management represents a segment of the *operational-bureaucratic apparatus*. For the macro variable "*instruction and education*" the ministry of education (which may be made up of specific ministries for schools, universities and higher education institutes) sets up management specialized according to each type of sub-variable. Increasingly more detailed local bureaucratic apparatuses are then put in place, down to the individual educational institute (education managers, headmasters, deans, etc.).

This arrangement means that no social system can control the homeostasis of the social variables which are vital for the maintenance of its own autopoiesis without a multitude of *Ring* holarchies operating daily in every part of the territory. This control of the social *macro* variables operates parallel to the mixture of *micro* controls on the behavior of single individuals; these controls derive in large part from the control of the macro variables.

It is not easy to examine the autopoietic network of the *Rings* operating at the various levels of the social system, given their ubiquity and extreme variety. Two difficulties are immediately apparent. Above all, control systems that operate in the socioeconomic system are usually *multi-lever* and *multi-objective*; the difficulty in recognizing them is due to the fact that, on the one hand, many of the control levers—for example, those that act on social decay and on the violence in our stadiums—are not known or immediately perceivable, and on the other that there are interesting, varying objectives over time that present sudden shifts linked, for example, to unexpected political changes, population flows between regions, natural and social calamities, and so on. As a result it is difficult to detect the deviations from the objectives and to identify the control variables for their elimination. A general model of such control systems could be represented as that in Fig. 7.15.

A second difficulty comes from the fact the socioeconomic system is made up of a network of subsystems of varying types and sizes which are closely interconnected, so that any control of a system influences the other systems it is linked to. For example, the control of labor productivity and employment levels—which are variables from the economic environment—implies the control of consumer optimism for the future, which is a typical variable of the social environment. The control of the level of national income implies control of consumption and investment at a national level; but this imposes a control at the regional level, and thus at the municipal level, all the way down to the company level. The control of employment implies both a production and administrative control policy. Controlling the spread of drugs implies control of the effectiveness of the educational system and of the quality of the family's educational activities.
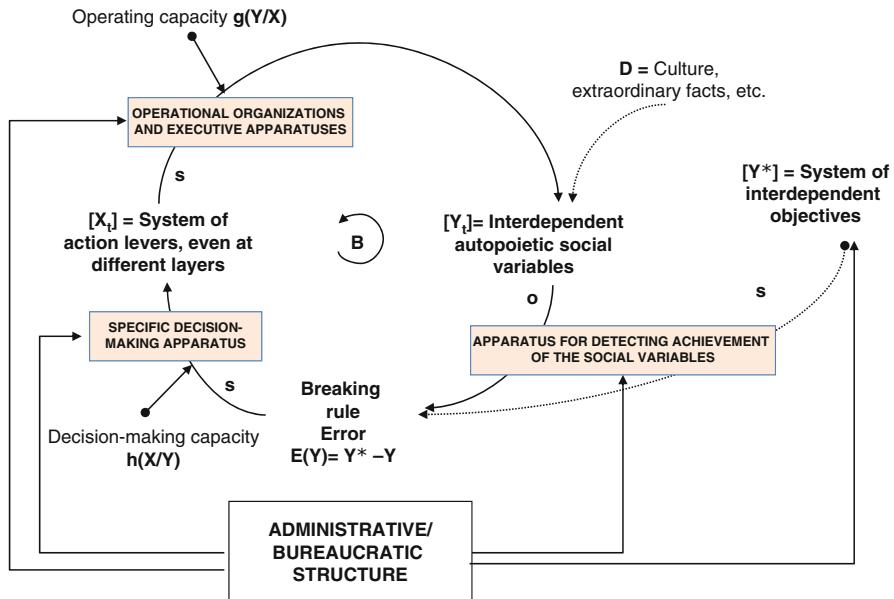
Fig. 7.15 General model of control systems for the autopoietic variables in social systems

7.7 *Rings* That Regulate Some Fundamental Variables in Social Systems

The preceding methodological observations, though brief, allow us to consider the typical control systems needed to regulate several macro variables fundamental for the preservation of social systems. I shall start with two particularly significant ones: the control of *traditions* and the control of *violence* (Mella 2012).

Surely no one will argue that the most deeply rooted tradition in any social system is the learning and maintenance of the *mother tongue* by the inhabitants of an area of varying size. This is a very powerful control system given that, as reported by SIL International the Ethnologue (2013) site, there are at present 7,105 languages, of which 2,304 are in Asia and 2,146 in Africa (Lewis et al. 2013, on line). The maintenance of spoken language—together with the written one, with its more narrowly codified syntax, which is less affected by practice—among the members of a community thus appears as one of the most powerful control system operating in human society. The linguistic control system teaches not only the rules of the new language but also the linguistic and dialectic inflections, setting off the levers to maintain them indefinitely by trying to use the instruction levers to eliminate speaker s (how often have scholars despaired over the many s in syntax and orthography!), according to the general model in Fig. 7.16.

At times the family of the newborn knows only one language, which is different from those of the other members of the community (immigration, deportation, etc.);
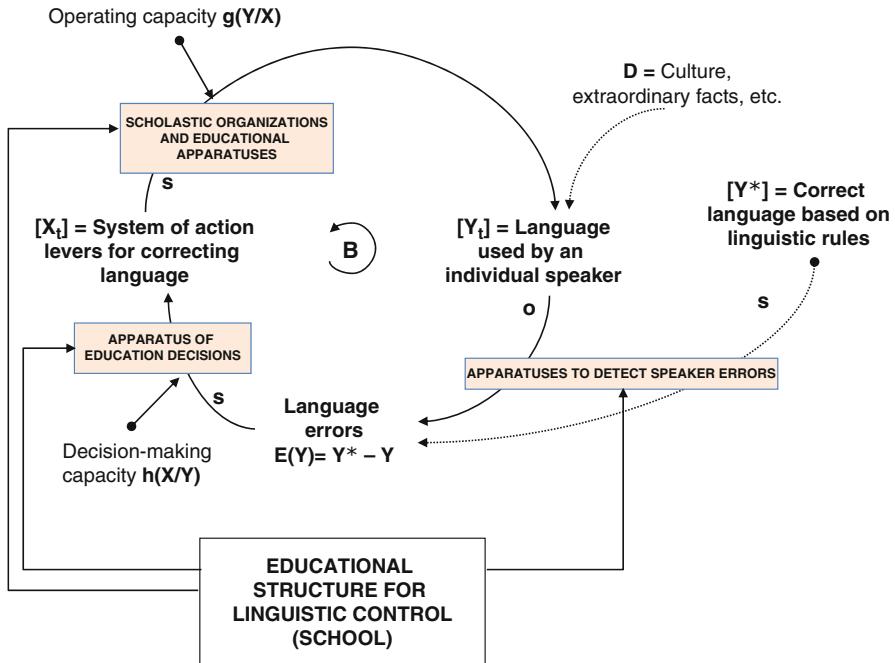
Fig. 7.16 *Ring* to control and preserve languages

in this case, this language is passed on as the mother tongue of the newborn, but the linguistic control system acts to eliminate the language gap between the newborn and the language normally spoken by the other members of the society, as the newborn is forced to learn a second language that is dictated by the educational control system of the community. Usually this second language is held to be rewarding, and is thus passed on to the offspring (dialects are spoken today only by the elderly, for whom they represent the mother tongue, since the youth are taught the national language). At other times, in the same social environment several communities coexist, each characterized by its own language; in this case there are as many control systems as there are different languages. If the languages are equally important with regard to interpersonal relations, and if marriages between people with different mother tongues are frequent, then the parents may pass on different mother tongues to their children at the same time (e.g., Switzerland). I will deal with this kind of control system in the Sect. 7.8 below, since the maintenance of the mother tongue is carried out by means of a combinatory system that operates automatically.

Violence is deeply rooted in our competitive nature and often degenerates into highly destructive behavior. Precisely because violence represents behavior inherent to man and is manifested in different contexts (family, work, on the roads, in stadiums, etc.) and with differing intensities (verbal or physical violence, with or without weapons, robberies, thefts, etc.), the social system must keep the level of

violence under the tolerance threshold—which is the control objective—by using various multi-lever control systems acting at various levels through various types of levers: a system of prevention through education and job promotion, a system of punishment and repression by means of restrictive measures for violent individuals, the arrest of promoters of violence, the loss of civil rights, etc. Though apparently different in nature, organized crime, common delinquency, stadium violence, bullying at school, hazing in the barracks, mobbing at work, and so on, are all encompassed in such a control framework.

More complicated is the violence that takes the form of wars and ethnic conflicts of varying magnitude. Man is certainly not the only social being that engages in conflicts.

> One of the behaviors of ants most commonly observed in nature is overt aggression, often resulting in injury, death, and displacement of one colony by another. Auguste Forel was correct in observing that "the greatest enemies of ants are other ants, just as the greatest enemies of men are other men" [...] Territorial fighting is very general in the ants, sometimes widespread over the foraging areas and sometimes confined to the immediate vicinity of the nest. (Hölldobler and Wilson 1990, p. 398)

Today there are many ongoing conflicts in the world. Wikipedia (2013) lists more than 20 current major wars and conflicts, as well as about 20 minor wars and conflicts.

Even if man, as a belligerent being, displays apparently unpredictable behavior, there is always some control system that stops wars from going on for too long, leading to the annihilation of one of the warring sides. This is testified to by the fact wars end sooner or later (even the Hundred Years' War only refers to a limited historical period). The main difficulty in identifying such control systems is identifying the constraints that oblige the belligerent parties to sit down at the peace table.

The so-called "religious" wars persist for a long time precisely because religions are deeply rooted in man's soul, especially in terms of prejudice and hate, rather than existing at an exclusively rational level. The outcomes of such wars are often horrific and their length unpredictable. Thus, the constraint that the war must end with the complete elimination of the "enemy," even if, rationally speaking, this appears too extreme, has occurred on many occasions. Other forms of wars of "conquest and defense" are motivated by the need for expansion (preservation) of the territory in which the social systems carry out their autopoiesis, due to the resources the territory offers as well as the need to subject the populations inhabiting it. In this form of conflict, while there are advantages that come with winning, there is still considerable damage even for the belligerent victor. A plausible limit to the continuation of war is linked to the trade-off between advantages and disadvantages. When the disadvantages exceed the advantages, it is highly likely, unless there are other noneconomic motives pushing the sides to fight "to the last man," that the war will end due to discouragement on the part of one or both of the belligerents, similar to Richardson's model on the limits to arms growth (Sect. 1.6.4). More realistic constraints that tend to limit the damage from carrying on a war are the reduction in the enemy's offensive capabilities below a given

threshold, thus minimizing the fear of the enemy, or a reduction in the defensive capabilities below limits perceived as dangerous for survival, so as to provide incentives to the enemy to take action to search for peace.

Two other fundamental control systems for autopoiesis in social systems are the control of *health* and *births*. The objective of the former is to increase the percentage of individuals who reach the limit of *life expectancy* at birth, that is, the average number of years an individual, from birth, can expect to live within a certain population. The system also aims to increase life expectancy itself. Since *life expectancy* depends on economic and cultural factors in the population, on the state of society, and on environmental conditions, the control system for this variable uses a certain number of control levers, among which fostering progress in medicine, improving physical hygiene, preventing disease through a system of early detection of infections and epidemics, diffusing dietary education, improving work conditions, and preventing and reducing the infant mortality rate due to work-related reasons or civil violence.

Births help maintain the population; however, too high or too low a birth rate with respect to the limit or objective necessary or useful for maintaining autopoiesis—also taking into account the death rate—necessitates a control system with multiple levers, which is not always easy to identify or hypothesize, since we do not know the optimal growth rate of the population. Births depend on psychological, economic, and social conditions whose effects are not always known; thus, the activation of appropriate control levers is often difficult. Such levers will impact lifestyles and work habits, family culture, the need for free-time, assistance for mothers and children, the condition of schools, and sports activities, requiring complex, long-term, and costly policies by the governance. Often favorable life conditions slow down births, while dangerous situations accelerate them. In some cases the existence of appropriate facilities for the care of children (pediatric clinics, nursery schools, kinder-gardens, summer camps, sports centers, and so on) are useful, while in others they have no effect.

As anticipated in Sect. 7.2, in some social systems during given historical periods there has been an attempt to control an excessive birth rate through restrictive measures sanctioned by law (prohibiting couples from having more than one child), or through economic levers when the birth rate is too low (aid for large families, tax breaks, etc.). Sometimes references to patriotic values can favor births; at other times such values are not effective. In short, every social environment must necessarily activate a hierarchy of coordinated control systems to control the birth and death rates, but the observation and planning of adequate control systems is anything but easy, as shown by recent history and daily news stories.

I will only briefly mention other important intentional control systems present in every socioeconomic environment.

The autopoietic process in a social system depends on the fundamental control systems of *nutrition* or malnutrition in the population, as these influence the birth and death rates, work capacity, and the levels of defense against epidemic diseases. It is not always easy to design effective *Rings*, as testified to by the presence of endemic malnutrition, diseases and infant death rates that are difficult to eliminate

(sleeping sickness, leprosy, etc.) or by the recurring outbreak of flu epidemics. Normally the nutritional level of the population is a *symptom* whose elimination is almost always achieved through simple symptomatic or extraordinary levers (humanitarian aid, dietary help); it is more difficult to identify the *structural levers* that can reduce the variance between an acceptable level of dietary wellbeing and the effective level that falls below this standard.

To conclude, I will mention several *Rings* that involve the economic variables, even when the economic system is so bound up in the social system that social variables are closely intertwined with economic ones.

The trends in economic variables (consumption, production, savings, investment, loans, exports, imports, price and interest levels, the levels of transfers and taxes, etc.) representing the entire economy of the social system are determined by the individual behavior of productive and financial organizations and by consumers, which as a whole determine the level of the nation's wealth, which is related to the level of national income and the gross domestic product, which in turn depend on the level of production, consumption and prices, and thus on investment, production costs and capital costs, in a typical dynamic system. In order to control economic behavior, which at the aggregate level produces global economic flows, the governance of the socioeconomic system introduces a number of forms of *regulation*.

Regulation—that is, government intervention in the market—is the process whereby the governance of the numerous economic *Rings* (activated by the economic control structures: ministries and other subdivisions of authority) identify the values to assign to the fundamental *control levers* in order to permit the *economic variables* to achieve the system's objectives or to bring these values back within the appropriate limits or respect the constraints of the economic policy the governance (government) feels appropriate for promoting the economic autopoiesis lying at the base of social autopoiesis.

> There are at least two large chunks of the economy that the competitive market model obviously does not describe or even purport to describe. These are the huge and growing public sector, the allocation of resources which is determined not by the autonomous market but by political decisions, and the public utilities [...].
>
> To be sure, the government influences the functioning of the private, competitive sectors of the economy as well in many ways—for example, by regulating the supply and availability of money, enforcing contracts, protecting property, providing subsidies or tariff protection, prohibiting unfair competition, providing market information, imposing standards for packaging and product content, and insisting on the right of employees to join unions and bargain collectively. In principle, these influences, however pervasive, are intended to operate essentially at the periphery of the markets affected. . . .In these sectors the government does not, or is not supposed to, decide what should be produced and how or by whom; it does not fix prices itself, nor does it control investment or entry on the basis of its own calculations of how much is economically desirable [. . .] In contrast, the government does do all these things with the public Utilities. Here the primary guarantor of acceptable performance is conceived to be (whatever it is in truth) not competition or self-restraint but direct governmental prescription of major aspects of their structure and economie performance. . .control of entry, price fixing, prescription of quality and conditions of service, and the imposition of an obligation to serve all applicants under reasonable conditions. (Kahn 1988, p. 2/1)

I shall only briefly expand on the above discussion by observing that ensuring the growth rate objectives are achieved is the task of government economic policies; however, the widespread and continuing difficulties in producing the desired growth in national income and GDP testify to how difficult it is to identify and regulate the necessary control levers. Together with the growth in national income and GDP, *inflation* must also be kept under control. The fact there is still inflation in many countries and that governments are constantly concerned it may rise again is the clearest evidence of the difficulty in planning an effective control system. It must be noted that the difficulties in controlling inflation also depend on the fact that, contrary to national income and GDP (which are variables that are to be controlled), *inflation is a symptom* of the lack of control of the socioeconomic system's most important variables—costs of particular factors (gold, oil, copper, iron scrap, etc.), production costs, productivity, etc.—whose trends depend on other variables which are not always economic in nature: the level of social and organizational conflict, confidence in the future, demand levels, the social security system, the propensity to save and to invest, and so on. In fact, in order to keep inflation under control we must identify both the variables which are truly relevant in influencing inflation and the levers that can be adopted to produce values compatible with economic growth at low levels of price increase.

I will just mention in passing the control systems for unemployment, productivity, and tax evasion, systems which are extremely important for every country's economic growth but which are difficult to design, since the dynamics of the variables to control depend on social and economic control levers which are not easy to identify.

7.8 *Rings* Within Collectivities as Combinatory Systems

A particular aspect that makes observing and designing control systems interesting but problematic is the fact the socioeconomic environment is made up not only of networks of systems and subsystems, which were discussed in the preceding *section*, but also of collectivities that I have defined as *simple social systems*—or simplex systems (Mella 2005a, 2008)—or, more generally, as combinatory systems.

A combinatory system is a simple social system formed by a group of similar individuals, or agents (humans, animals, or other living units), forming the *base* of the system, not organized by hierarchical relations or interconnected by a network or tree relationship, that produces an analogous micro behavior (pedestrians walk along the sidewalk, elephants in the savanna) that can produce some observable micro effects (individual pedestrians leave garbage behind; individual elephants head towards other elephants). Combined together, the micro behaviors produce a macro behavior (general tendency to litter; a compact herd forms), which can produce an observable macro effect (a garbage pile forms; the herd opens a trail in the forest), which, in turn, conditions the subsequent micro behavior of the

individuals (the garbage pile represents information about where one can leave garbage with impunity, the herd gives information that there is protection available).

Thus, on the one hand, the macro behavior of the system, as a collective unit, derives from the *combination* (appropriately defined) of the analogous micro behaviors of the individuals (from which the name combinatory systems derives); on the other hand, the macro behavior determines, conditions or guides the subsequent micro behaviors. This reciprocal relation can be defined as micro–macro feedback. I assume that a *necessary* and *sufficient* condition for a collectivity (observable or assumed) to be considered a combinatory system is the existence of a feedback between the *micro* behavior of the individuals and the *macro* behavior of the collectivity constituting the system. Combinatory systems are very common in the biological and social worlds, but we are not used to perceiving them (Mella 2005a). Examples of these are the systems that lead to the formation of groups of various objects, from cities to industrial districts (Mella, 2006), or to the spread and maintenance of particular characteristics, from language to traditions, fashion to collective states of mind.

The action of combinatory systems occurs because the macro behavior—or the associated macro effect—translates into *global information* (of varying types) that represents a *global objective* individuals must or desire to achieve; this objective guides the choices of individuals who, noting a distance between their own micro behavior or micro state and that of the overall state of the system, try to eliminate the gap by adjusting their individual and collective behavior. The system's macro behavior, or macro effect, represents (or conditions) the *objectives* behind the individual choices.

It clearly follows from this that combinatory systems function due to the presence of *microRings* which, operating at the individual level, lead to uniform micro behavior by individuals in order to eliminate the (gap) with respect to the objective that is represented—or revealed—by the global information (macro behavior or effect). The above description can be better understood through the model in Fig. 7.17.

For this reason combinatory systems can also be called *self-produced global information systems*, in order to distinguish them from *local information systems*, whose typical model is represented by complex systems simulated by cellular automatons. When the actions of individuals are simultaneous, the agents of the system appear to *synchronize* their micro behavior. However, each micro behavior updates the global information, and this recreates a divergence that exerts even more influence on the individuals to conform to this information.

The mutual dependence between the *micro* and *macro* behavior (or their states or effects) represents the *micro–macro feedback* [**R**], which is maintained thanks to the individual control systems represented by the loops [**B**] in the lower portion of the model in Fig. 7.17, so that the combinatory system in principle repeats its cycles until the individual actions can update the self-produced global information which makes clear to each individual the distance, E(me), of the individual state (me) with respect to the state representing the objective (me*).
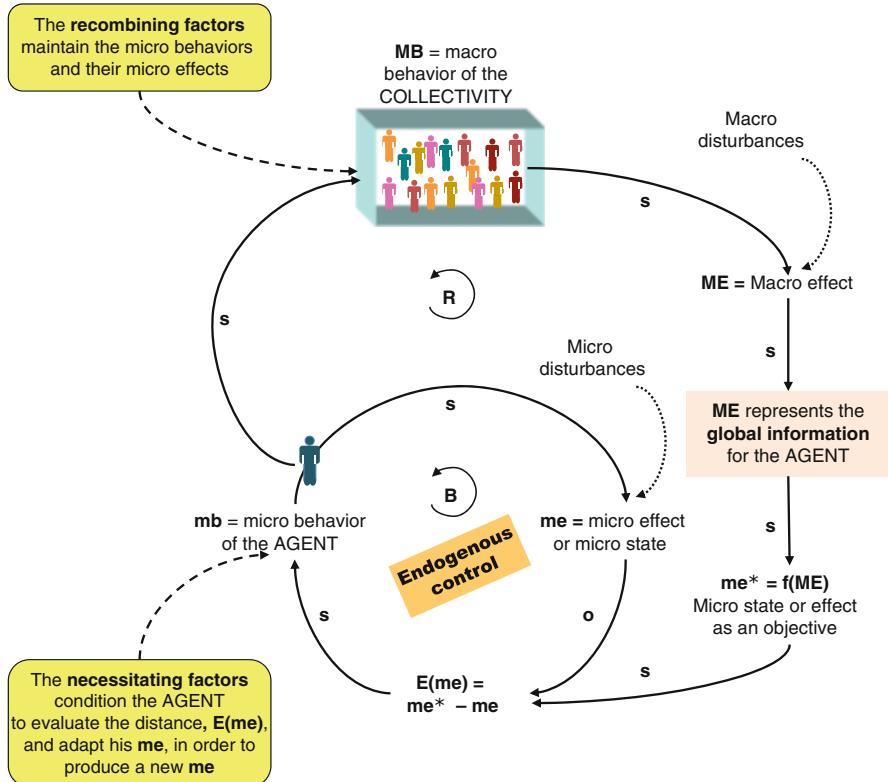
Fig. 7.17   Schematic general model of a combinatory system

The feedback arises and is maintained by a set of *necessitating factors* which force the agents to adapt their micro behavior to the system's macro behavior or effect; the feedback is also maintained by the action of a set of *recombining factors*, which lead the collectivity as a whole to recombine the micro behaviors, or the micro effects, in order to produce and maintain the macro behavior, or the macro effect (see Sect. 7.9). Recognizing the existence of a *micro–macro feedback* and understanding the nature of both the *necessitating* factors and the *recombining* ones is indispensable for interpreting collective phenomena as deriving from a combinatory system.

The logic of combinatory systems allows us to identify a wide range of phenomena that depend on the individual control systems that lead to micro behavior compatible with the global information (macro behavior or macro effect) produced by collective action. To more easily understand the *modus operandi* of combinatory systems and the role of global information, let us consider the simple phenomenon, familiar to all of us, of a "*buzzing noise* arising in a crowded room."

What causes this buzz to form? It arises from the voice level (micro effect) of those present when they speak to each other (micro behavior), which is recombined by the shape of the room. But why do those present speak in a loud voice (macro behavior)? Because there is a buzzing noise (global information) which prevents them from being heard; this becomes the minimum constraint (objective) to exceed in order to be heard; if the voice level is not adequate (gap) it is not possible to communicate. Thus, if the buzz increases, those present, in order to be heard, must raise their voices. It seems they do this all together, as if the global information forces them to synchronize their micro behavior; but this causes the buzz to increase further, which obliges those present to raise their voices even more, which increases the buzz, which forces those present... etc., as part of a reinforcing loop that takes the buzz to the maximum bearable level (constraint); once this is reached the individual control systems induce the speakers to be quiet, as we have all witnessed first-hand. The next *section* will provide a formal representation of the combinatory system I have described; for now I will present the system that produces the process for "passing on a *language* within a population," which has briefly been dealt with in Sect. 7.7.

Let us begin by stating that all parents pass on their mother tongue to their children (micro behavior), and that the children learn this language (micro effect). The population communicates (macro behavior) using the mother tongue (macro effect), which represents the global information that obliges families to teach that language to their children (constraint/objective) in order not to disadvantage them in their communication (s). The feedback is evident but—remembering the bad marks inflicted on us by our teachers because of syntax s—we can imagine there are also external control systems that both reinforce the main combinatory system, by detecting deviations between the language of the group members and the codified mother tongue (Eccles and Robinson 1984, p. 141), and try to eliminate these by using the traditional levers of academic teaching to correct the syntax s. In Italy the Accademia della Crusca (National Academy for the Study and Preservation of the Italian Language) is the depository of "proper speech" in the Italian language. The existence of so many languages and equally numerous dialects show how powerful this combinatory system is and how efficient the individual and external control systems are. Thus, combinatory systems for the spoken language—to the same extent as those for the written language, with its more strictly codified syntax that is less affected by accepted practice (consider all the syntax and spelling s that have driven scholars crazy!)—are some of the most powerful systems operating in human society.

The process behind a "*fashion* spreading" appears different from that regarding the maintenance of a mother tongue; but the former is also produced by the action of a combinatory system and by the individual controls this entails. How do fashions arise, how do they cease, and why do some return while others die away forever? A fashion—whether clothing, watches, cars, toys, the use of a particular linguistic form, and so on—arises in a given environment through a novelty introduced (micro behavior) "by chance" by a given creator for the purpose of distancing himself from custom (objective to exceed). If the novelty is not pleasing, costs too

much or has other disadvantages, the episode will remain an isolated one and no combinatory system will arise. If a sufficient number of individuals (minimum activation number) see the imitation of the innovation as an objective to achieve and follow the novelty in order to eliminate the deviation from the new desired state, then that innovation becomes "in fashion" and its spread represents the global information of the new "tendency." As a result a macro behavior is produced that refers to the collective system, which leads to an increase in the desire to imitate, thereby causing individuals, conditioned by the individual control systems, to undertake micro behaviors that imitate the novelty, with the objective of eliminating the between the desired and normal states. The resulting micro effects consist in the purchase of goods which are the object of imitation, a further increase in the desire to imitate, and the resulting macro effect of the further spread of the good that is "in fashion."

Similar to combinatory systems for the spread of a fashion are those that move individuals to "pursue *records* of whatever kind." Even if Pierre De Coubertin stated that "*it is important to participate, not to win*," those who compete are not content merely to equal the record (objective to achieve), sparing no effort to eliminate the gap and exceed it, since the record, in any sporting event, is the global information that identifies the "absolute best." When the effort comes from a multitude of competitors the record gradually improves, thus raising the level of the objective the individual competitors must achieve. It is easy to see in this mechanism the continuous action of individual control systems to eliminate the deviation between the record—which represents a dynamic objective—and the actual performance of the competitor, who represents the ARCHETYPE of "Strengthening Goals" (Fig. 5.12). It is normal today in national competitions for even the last-place runner in the 100-m dash to break the record from the 1896 Olympics, when the gold-medalist, Thomas Burke (USA), won with a time of 12 s; the current women's world record is held by Florence Griffith (USA): 10.49 s in 1988. The tributes for the record holder, the considerable prize money and social gratification are levers used by the collectivity to push competitors to improve their performance by activating control systems that, through the levers of training, improvement, the learning of techniques and tricks—and often the use of banned substances as well—seek to eliminate the gap between individual performance and the objective, despite accidents, victims and physical damage. Here, too, it is not difficult to see how this combinatory system is, in turn, reinforced by environmental control systems such as: the spread of sports education, government sports policies, and the building of more gyms, pools, and other sports facilities.

Let us move to another phenomenon, which is completely different but still produced by a combinatory system. Those who often travel on highways know how annoying and dangerous "*wheel* ruts" are; these are the parallel ruts that form on the pavement from the constant passage of heavy trailer trucks over the same trajectory. The trucks' wheels continually use the same section of pavement, especially on hot days when the pavement is softer, causing a typical subsidence that becomes a long wheel rut if observed over a long stretch on the pavement mantle. Wheel ruts occur especially on long straightaway sections of narrow lanes since, in order to stay

within their lane and avoid invading the adjacent one, the drivers of large vehicles stick to the same strip of highway mantle. Ruts are less common along curves, since it is harder to keep the same trajectory in this case. Once formed, wheel ruts force trailer truck drivers to drive within the ruts on the highway so as to avoid dangerous swerving and route adjustments that could cause them to invade the adjacent lane. We can clearly see the action of the individual control system of the single drivers that try to eliminate s with respect to the trajectory imposed by the wheel ruts, which represent the global information that forces them to adopt a uniform micro behavior.

These few examples show that a large number of socioeconomic and sociopolitical phenomena can be convincingly explained using the model for combinatory systems. Without attempting a complete description, but merely to make order out of the many social phenomena produced by combinatory systems, I shall now present a typology composed of five classes, based on the macro effect produced.

1. *Systems of accumulation*, whose macro behavior leads to a macro effect which is perceived as the accumulation of objects, behaviors, or effects of some kind; this logic applies to quite a diverse range of social phenomena, among which the formation of urban or industrial settlements of the same kind, of industrial districts, the accumulation of garbage, graffiti, and writings on walls, and the clustering of stores of the same type in the same street (Mella 2006).
2. *Systems of diffusion*, whose macro effect is the diffusion of a trait or particularity, or of a "state," from a limited number to a higher number of agents of the system; systems of diffusion explain quite a diverse range of social phenomena: from the spread of a fashion to that of epidemics and drugs; from the appearance of monuments of the same type in the same place (the towers of Pavia, for example) to the spread and maintenance of a mother tongue or of customs.
3. *Systems of pursuit*, which produce a gradual shifting of the system toward a *dynamic objective*, as if the system, as a single entity, were pursuing a goal or trying to move toward increasingly more advanced states; this model can represent quite a different array of social behaviors: from the pursuit of records of all kinds to the formation of a buzzing in crowded locales; from the start of feuds and tribal wars in all ages to the overcoming of various types of limits.
4. *Systems of order*, which produce a macro behavior, or a macro effect, perceived as the attainment and maintenance of an ordered arrangement among the agents that form the system. Systems of order can be used to interpret a large number of social phenomena: from the spontaneous formation of ordered dynamics (to an observer) in crowded places (dance halls, pools, city streets, etc.) to that of groups that proceed in a united manner (herds in flight, flocks of birds, crowds, etc.); from the creation of paths in fields, wheel-ruts on paved roads and successions of holes in unpaved roads, to the ordered, and often artificial, arrangement of individuals (stadium wave, Can-Can dancers, Macedonian phalanx). Systems of order can also explain populations of insects, typically ants, which act by creating an "aromatic potential field" by spreading *pheromones* or other permanent messages; the increasing concentration of pheromone (global or

macro information) increases the probability that each agent will move in the direction of that site. The micro–macro feedback is quite evident.

5. *Systems of improvement and progress*, whose effect is to produce progress, understood as an improvement in the overall state of a collectivity normally conceived of as a consequence of evolution.

Schematic and heuristic models of such systems are presented in Sect. 7.13.1.

7.9   The Control of Combinatory Systems

As shown in Fig. 7.17, combinatory systems can operate thanks to the micro control systems that allow agents to adapt their individual states to the macro state of the system as a whole. Nevertheless, combinatory systems can, in turn, also be controlled to favor or inhibit the attainment of certain individual and collective results considered to be harmful or useful.

Harmful behavior could be one that leads to the writing of "graffiti on a wall"; every instance of graffiti (micro effect) left on the wall (micro behavior) increases the mass of graffiti (macro effect), which in turn increases the probability that individual passersby will leave new graffiti, in a typical micro–macro feedback process. "Garbage piles" also influence passersby who want to get rid of their garbage; but each time garbage is left, the pile increases not only in size but also regarding its "attracting power" (global information), thereby producing an undesirable effect. Without an external control on the combinatory system, the effects of collective action based on self-produced global information could lead to harmful, even disastrous effects.

When, on the contrary, the combinatory system produces a useful macro effect, then the external control could accelerate the formation of the macro effect; we see this clearly, for example, in the formation of industrial districts (macro effect), which are formed when new firms are enticed to locate in a particular district (micro behavior), which makes the area even more advantageous by offering more employment and increasing the local wealth.

What happens if the combinatory system acts freely, without any external control? In some cases a type of self-control is produced: no one can continue raising his voice to overcome the buzz in the crowded room; such a volume is reached that people spontaneously stop talking, and then the buzz diminishes (only to soon begin to rise again). In other cases the combinatory system finishes operating when the maximum admissible density (or maximum level of the macro effect) is reached, making any further individual action impossible.

In medieval Pavia (northern Italy) a phenomenon occurred that was unique in the world. The "tower phenomenon" began around 1000, perhaps some decades before, and rapidly developed in 1100 and the following century, so that by around the year 1300 many had already been ruined. In 1570 the historian Breventano (1570) mentioned more than 170 towers, Spelta (1602) around 100, and Zuradelli (1888)

counted 76, a very impressive number if we consider that a recent map of the town reveals 71 traceable towers within its walls, 7 of them almost intact. The towers had no useful function; they only reflected the desire of the noble families of Pavia to display their wealth. They were the visible icons attesting to the fact the family had attained wealth and could allow itself to make a show of this by using this symbol. I have carefully studied this by creating simulation software that provides results that conform to the historical data (Mella 2013). The appearance of a swarm of towers is certainly the macro effect of a combinatory system that is easy to understand. The towers represent the global information that many families have achieved wealth. In order not to be any less appreciated, new families built (micro behavior) their own tower (micro effect), thereby making the global information more significant; the denser the "swarm of towers" became, the more those families without a tower felt inferior and were spurred to build their own tower.

Why did this phenomenon die out? The plausible explanation is the following: during that historical period the area inside the town's defensive walls extended over a surface area of about 900×900 m. If we consider that the towers were built in the corner of palaces belonging to noble and wealthy families, and if we assume that on average a palace was 30×30 m (which can be objectively verified), then there were no more than 300 palaces for rich and noble families in Pavia, if we take into account the surface area occupied by the less well-off townsfolk, as well as by the town's squares, roads, churches and public buildings. We thus have a possible explanation: the towers had become so numerous (a density of one tower for every two palaces) that they ended up losing their power of attraction; the global information had changed from "if you don't have a tower you're not considered rich" to "if you build a new tower you're a parvenu." The phenomenon had come to an end.

Leaving aside the spontaneous cessation of combinatory systems, there is no doubt that very often *external controls* are carried out on combinatory systems using external control systems constructed expressly for that purpose. In order to understand how these external control systems for combinatory systems operate, we must recognize that the micro–macro feedback that produces the micro and macro behavior requires the contemporaneous presence of *necessitating* and *recombining* factors.

I define a *necessitating factor* as any element—a constraint, rule, condition, law, conviction, etc.—that "obliges" each individual in the collectivity to adapt his micro behavior to the macro behavior of the system by activating the individual control system that detects the gap between the individual's micro state and the macro state of the collectivity, leading to the activation of the control levers needed to eliminate this gap ($E(Y)$). The stronger the *necessitating* factor is, the more the micro behavior is activated. In combinatory systems made up of persons, the necessitating factors often originate from conscious motivations: necessity, convenience, opportunity, the desire not to be inferior, and so on. At other times these factors can also be "natural" and act unconsciously, since they derive from the genetic or operational program of the individuals that form the *base* of the combinatory system.

To understand how *necessitating* factors operate, let us reflect on the highly powerful combinatory system involving the spread of drugs and ask ourselves what causes the addiction that moves the drug taker to buy and sell those substances (conscious need), thereby producing the collateral macro effects represented by the network of drug pushers and the cultivation and importation (macro effect) of such drugs. What moves the driver of a trailer truck to remain in his lane, if not the need to avoid dangerous swerving? What instincts incite individuals to perpetuate feuds and start wars?

The existence of one or more necessitating factors is indispensable though not sufficient; it is also necessary for the system—through some *recombining factor* (rule, convention, algorithm, nature of the environment the agents operate infect)—to be able to recombine the micro behaviors (or micro effects) to produce the macro behavior (or macro effect) which, through the *micro–macro feedback*, is able to act on the necessitating factors. The recombining factors can also operate in various ways.

Consider how different the recombining effect is regarding the formation of the buzzing noise when the room is large as opposed to high-ceilinged or small; or flat as opposed to arched; or when there is a small or a large number of speakers. In the countryside it is much simpler to scatter one's refuse haphazardly than to accumulate it. In cities the systematic removal of garbage makes it almost impossible for the combinatory system that produces garbage piles to start up. In an area with a high population density it is easier for a fashion to spread; even the action of advertising promotions, together with the abundance of sales outlets, makes the recombining of this system much stronger, faster, and more persistent.

From the above comments it is easy to see that when the *macro* behavior or the *macro* effect of the combinatory system (ME_t) must be guided toward a desired objective or limit (ME^*), set by some policy maker in the socioeconomic environment within which the combinatory system operates, specific external control systems can be activated whose control levers, $[Xn_t]$, are represented by *reinforcing* or *weakening measures* (actions, provisions, constraints, limits, obligations, etc.) which, operating at a macro and/or micro level, modify the *recombining* and *necessitating* factors, influencing the *macro* and/or *micro* behaviors, thus directing the *macro* behavior of the combinatory system, as shown in Fig. 7.18.

As an example, let us apply all this to the combinatory system of *accumulation* that gives rise to industrial settlements or the formation of industrial districts (Mella 2006). When a certain area offers a positive differential in terms of economic efficiency compared to others (necessitating factor), then a certain number of entrepreneurs may decide to locate there (micro behavior) and set up production and commercial units (micro effect). The presence of a group of firms always produces economic advantages for future settlements (recombining factor), and this attracts new entrepreneurs that weigh these advantages against the disadvantages of locating in nonindustrialized areas, in so doing activating a typical micro–macro feedback process. When the *minimum activation density* is reached (number of firms needed to produce a positive differential in economic efficiency) the combinatory system starts up and, in subsequent cycles, produces the collective
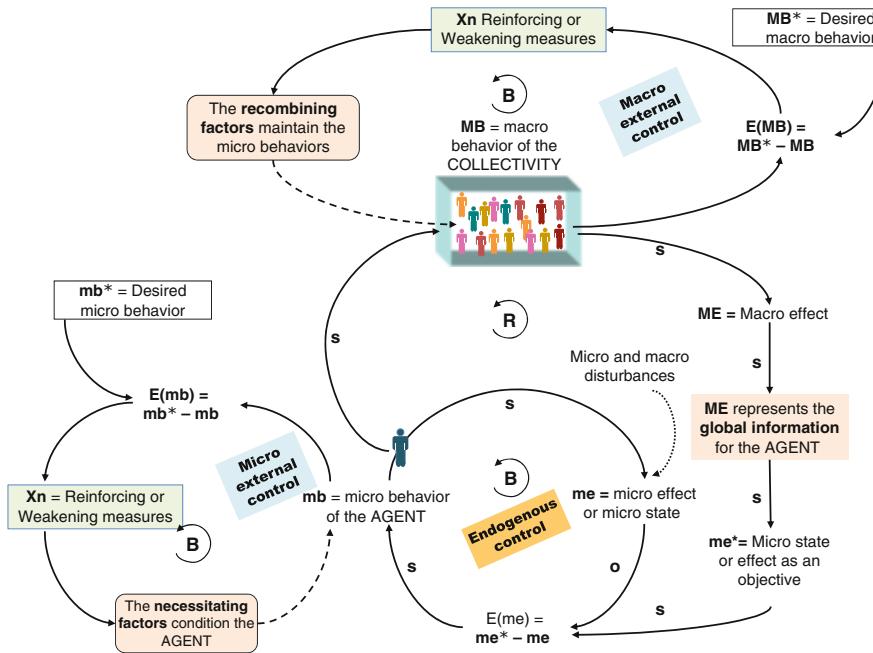
Fig. 7.18 External macro and micro controls in combinatory systems

inflow of other production units (macro behavior) and a rise and expansion in the typical industrial and commercial districts (macro effect). The system can also be controlled by both weakening actions (urban regulations, higher taxes and charges, etc.) and reinforcing actions (outright grants to locate in an area, the creation of infrastructures, etc.), which create new economic efficiency differentials. Reinforcements and weakenings are the levers $[Xn_t]$ that control the combinatory system, and these can be applied both directly to the *macro* behavior—we shall define this as the *macro* control—or to the *micro* behavior, in which case the control will be defined as *micro* control.

The model of the combinatory system combined with the external control systems is shown in Fig. 7.18, in which we immediately see that the macro-level control is achieved through reinforcements and weakenings that act on the recombining factors, while the micro control requires reinforcements and weakenings to modify some necessitating factor. It should be clear that these are *external* control systems run by managers that are external to the combinatory system; these systems should not be confused with *internal* control systems, which are necessary to maintain the micro–macro feedback (Sect. 7.8). Also note that the micro control must not be confused with the *feedback* between the macro and micro behavior internal to the combinatory system.

I wish to proceed with the example of the productive settlements that form an industrial district or commercial zone, considering, however, how this phenomenon

can be produced by a combinatory system of *diffusion* that operates in the socioeconomic environment to favor the *endogenous* spread of new firms.

Some successful firms that locate in a particular area and that are able to internally develop their employees, managers, and professional staff may decide, after acquiring the necessary competencies and pushed by the desire for personal profits (necessitating factor), to undertake an activity to exploit these competencies (micro behavior), thereby creating new businesses (micro effect). If those who undertake an autonomous business venture are successful in their attempt, then the combinatory system can start up and the area will see the growth of new businesses (macro effect) that *endogenously* develop over the territory. The process spreads and, after subsequent cycles, the group of workers is gradually transformed into a group of businessmen (macro behavior), thereby creating industrial settlements (macro effect), which probably will require bringing in new employees from other areas.

Since this spread of businesses is considered to be the source of well-being for the entire area, various forms of *macro* and *micro controls* can be carried out by activating several levers such as: creating professional schools that guarantee initial job entry (macro control); the availability of risk capital and loans (micro control); incentives to start new businesses through facilitating regulations to favor, for example, young entrepreneurs (macro control); and the creation of safety nets for failure (micro control). The creation of businesses exogenously, or through public authorities, which possess the appropriate business characteristics (even small-sized firms, the need for small-scale collateral production, the development of competencies) can artificially start up the system, on condition that the critical mass necessary to produce the necessitating and recombining factors that produce the micro–macro feedback is quickly achieved.

7.10   The Tragedy of the Commons

Combinatory systems (Sect. 7.8) can produce *micro–macro feedback* with various consequences. In some cases the rationality of individual behavior is accompanied by a collective effect that leads to an improvement and to progress; in others to a negative, even catastrophic result. A particular combinatory system with *negative collective effects* is the one that leads to the exhaustion of scarce, nonrenewable resources that individuals can freely appropriate and use. This situation is normal whenever individuals, faced with a resource in danger of depletion, behave according to the ARCHETYPE of "short-term, local and individual preference" (Sect. 1.6.6 and Fig. 1.19), which dominates the prevailing form of individual–agent, leading the latter to adopt individual and rational micro behaviors that the combinatory system transforms into disastrous collective macro effects (two other examples of which will be presented in Sect. 7.13.3).
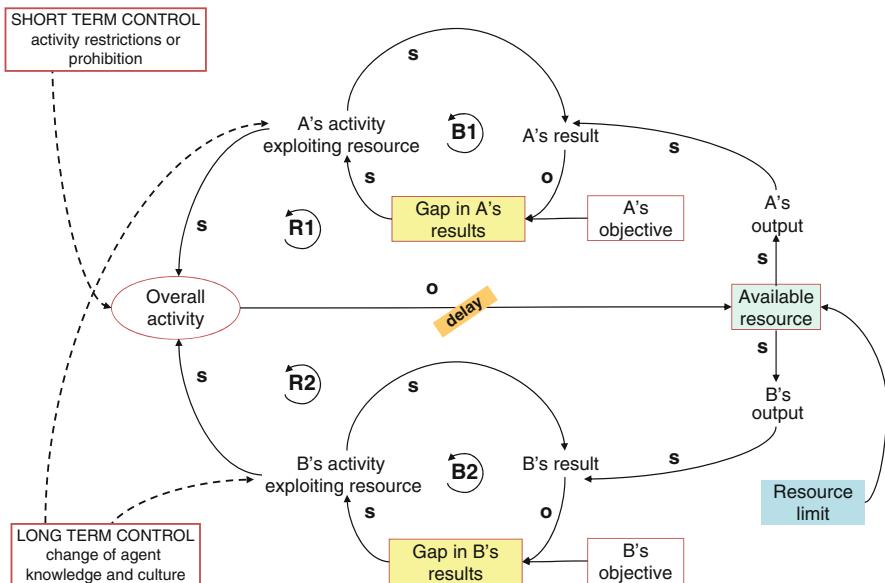
Fig. 7.19 Archetype of the Tragedy of the Common Resources (*Source*: adapted from Mella 2012, p. 251)

This model is so common in social systems that it is considered a true systemic archetype, known as the Tragedy of the Common Resources, or the Tragedy of the Commons, knowledge of which is useful for understanding the causes of many types of extinction or exhaustion and for finding remedies (Sect. 7.13.3). This combinatory system can also be activated in nonhuman populations competing for a territory or a food source; if not stopped through exogenous controls it can inevitably lead to the extinction of one of the populations concerned.

This combinatory system contains a certain number of individuals—the appropriators—who share a common resource available in limited quantities, called Common-Pool Resources, or simply "commons" (Hardin 1968, 1985; Ostrom 1990), which are not distributed based on some indicator of merit or need but freely exploited for individual activities, in order to achieve personal objectives whose outcome depends on the use of the resource. It is obvious that the more the resource is exploited individually the less is available collectively, so that even the activities that employ the resource are less efficient, and thus its use becomes increasingly "less economic." To achieve the objectives in a situation of lesser efficiency, the appropriators must intensify their activity, continuing to exploit the resource, no matter how scarce, until it is completely exhausted. The result is a *tragedy*, since those who exploit the resource are forced to interrupt or reduce their activity because of the exhaustion.

To better understand the action of this combinatory system it is useful to illustrate the simplified situation where there are only two agents, A and B, with external controls, as shown in Fig. 7.19, and where we assume the conclusions apply to any number of agents.

Technical Note. There are different ways of representing this combinatory system. I prefer the model in Fig.7.19 to that presented by Peter Senge (1990, p. 446), since the latter does not highlight the two individual control systems, that is, the fact that A and B have their own objectives even in the presence of diminishing returns from the use of the common resource. In the context of control systems, I feel that Senge's model might be completed as indicated in Fig. 7.19.

Referring to Fig. 7.19, the loops [**B1**] and [**B2**] are, in fact, individual control systems that show that the individual actions of A and B are undertaken to achieve a personal objective; if the common resource is abundant and the marginal performance is high, then these activities will produce a result in line with the objective. If there is no gap between the actual and desired result, A and B have no reason to intensify their resource-exploiting activities. Since we have assumed a limited availability of the common resource, the combined activities of A and B lead to a gradual reduction in the marginal performance of the resource, which is usually perceived with a (sometimes even considerable) delay.

Assuming they undertake the same activities, A and B see that their individual results clash and fail to achieve their objectives, which produces a distance, $E(Y)$, between expectations–objectives and results. This gap forces them to intensify their activities, which acts as a lever to achieve their expected outcomes. In any event, the efforts by A and B to pursue their objectives generate the reinforcing loops [**R1**] and [**R2**]—which correspond to the micro–macro feedback (divided into two loops to simplify the figure)—whose action inevitably leads to the exhaustion of the common resource, with negative effects for the individual and collective outcomes.

Figure 7.20 represents the whale hunting process (this example can be adapted to the hunting of elephants for ivory, seals for fur, etc.) under the assumption that the reproduction rate of the population of whales is lower than the hunting intensity, so that the simultaneous action of many whale hunters reduces the number of whales before they have had the chance to reproduce in adequate numbers, thereby diminishing the returns from hunting.

When the returns from hunting start to diminish, the whale hunters, in order to fill their oil and meat holds (the objective of every whale hunter), are pushed to intensify the hunt (remaining longer at sea), as shown in loops [**B1**] and [**B2**], until the whales reach the minimum admissible population, as shown in loops [**R1**] and [**R2**]. The two control systems, [**B1**] and [**B2**], can have different modes of action, all inevitably destined to result in the extinction of the scarce resource. The hunters could, for example, intensify their hunting activity not only by devoting more time to it but also by improving the quality of their equipment—radar, sonar, satellites, harpoon cannons, or teleguided missiles—with the result that the whales, no matter how rare they become, are inexorably identified and hunted. The result? There is no way out for the whales, unless controls external to the combinatory system intervene.

The *external* macro control at the collective behavior level could operate by implementing several levers; for example, a ban on whale hunting by world authorities, or a ban on the consumption of whale meat and fat by individual countries. The external micro control could operate, for example, by granting subsidies to convert
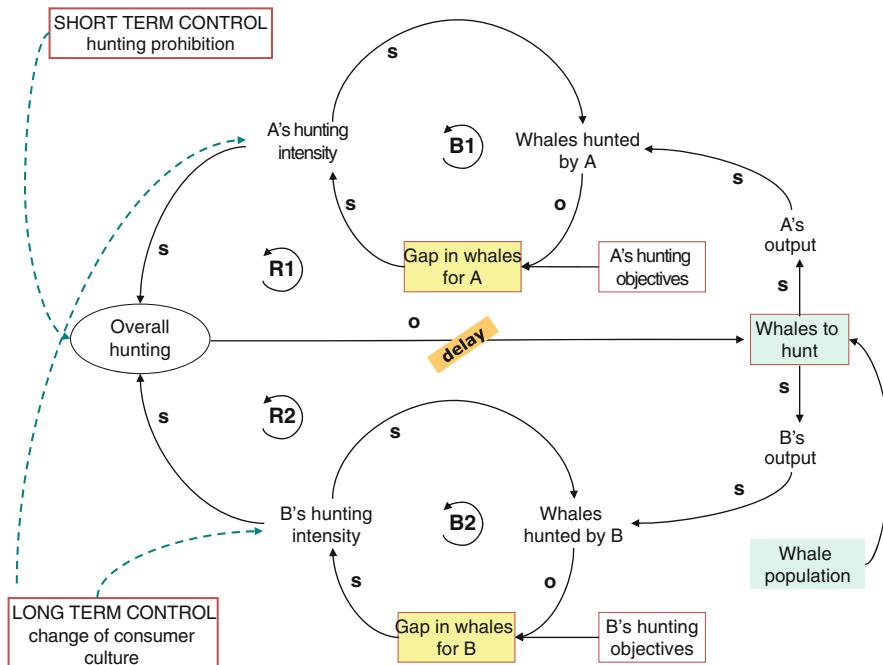
Fig. 7.20 Tragedy of the Common Resources in whale hunting (*Source*: Mella 2012, p. 253)

whaling ships to fishing vessels, or by initiating information campaigns to discourage consumers from eating whale meat, thereby making whale hunting less convenient from an economic point of view. In conclusion, in such combinatory systems individual interests guide the individual control systems that provide the impetus for the collective exploitation of scarce resources until they are depleted.

In The Tragedy of the Commons (1968), Garrett Hardin also emphasizes the fact that populations that share a common vital resource almost always head toward tragedy.

> Adding together the component partial Utilities, the rational herdsman concludes that the only sensible course for him to pursue is to add another animal to his herd. And another; and another. . . But this is the conclusion reached by each and every rational herdsman sharing a commons. Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit—in a world that is limited. Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons. Freedom in a commons brings ruin to all. (Hardin 1968, p. 1244)

If we observe the world around us we see this archetype in operation in many situations: too many fishermen fishing in waters with limited populations of fish; the search for oil, minerals and precious stones in the same area and, in general, the exploitation of limited natural resources by many users. We should not wonder if the tropical forests are reduced because their inhabitants try to transform them into arable land; if the aquifer levels in the desert fall because in many areas irrigation

systems are built that take water from them; if rivers run dry because alongside them ever larger areas are transformed into rice fields; if the water-bearing strata in the deserts are being lowered because in many places irrigation plants are being built that take water from those strata; or if rivers are drying up because along their banks larger and larger areas are being turned into rice fields. It appears that even Easter Island was abandoned after its inhabitants destroyed all the island's wood resources in order to build their Moai. This archetype also operates in the business field, since its effects are completely analogous to those produced by "arms" growth, where firms go "hunting" for consumers in a market heading toward the saturation point.

The presence of limited common resources can cause conflicts that represent a *second tragedy* included in the tragedy of the commons. The clearest evidence of this are the numerous wars fought over water and oil, which are among the most limited and scarce vital goods on earth.

How can the action of the archetype be countered? There is only one answer: the managers/decision-makers must externally control the combinatory system that produces the harmful effects that must be avoided by specifying several useful levers to reduce conflicts and manage scarcity, among which:

(a) delimiting the boundaries of the common resource, restricting and possibly limiting the appropriators;
(b) specifying the technologies that can be used to exploit the resource, placing restrictions and quantitative and time limits on its allocation;
(c) providing incentives for resource saving and, when possible, for its regeneration;
(d) setting up an outside authority to accurately and continuously monitor individuals to ensure they are respecting the rules for resource allocation and sanctioning violations of the rules for resource exploitation;
(e) discouraging conflicts between resource users and seeking to resolve conflicts that have broken out in a peaceful manner;
(f) encouraging collective participation in the choices regarding the above points in order to favor agreements on collective exploitation, self-discipline and shared social oversight.

7.11   *Rings* Operating in Social Systems as Complex Adaptive Systems (CAS)

Social systems, as autopoietic systems, must not control the homeostasis of the fundamental variables for the survival of the individuals and groups that make up the system, nor limit themselves to creating the *Rings* for achieving the autopoietic macro objectives based on the *policy* decided by the governance (Sect. 7.6). Social systems can survive for a long time only if they react and adapt successfully to the disturbances (disasters, wars, radical technological innovations, etc.) and threats their structure can undergo, which the individual's in the system must face.

Every structural disturbance influences above all the autopoietic organization of the social system, the interpersonal relations and the governance; this obliges the individuals themselves, the groups and the various coexisting autopoietic subsystems to react and adapt their mutual behavior and their relationships based on new forms of coexistence in order to survive, thereby also changing the political structure (rules and social control systems) and the economic structure (autopoietic objectives). The individual adaptive reactions lead to further changes in the social autopoietic structure which necessitates further individual adaptation, based on reciprocal control systems that also produce numerous cycles that end when the social system has absorbed the disturbances and found new homeostatic equilibriums between the individual behaviors and the basic variables for the system's preservation. These new equilibriums are maintained until new disturbances arise that require further adaptations.

Every social system, though composed of a number of agents with different individual features, social roles, economic motivations, different group memberships, and so on, represents a complex entity, of a *variable organizational nature*, capable of being "malleable" and thus adapting to disturbances (which do not overwhelm the system). In this regard it can be considered a CAS with the capacity for *structural adaptation* and *organizational evolution*, whereby it produces new policies and strategies to control the cohesion and homeostasis of the autopoietic variables, as shown in this concise definition:

> Definition (1): A CAS consists of inhomogeneous, interacting adaptive agents. Adaptive means capable of learning.
>
> Definition (2): An emergent property of a CAS is a property of the system as a whole which does not exist at the individual elements (agents) level. Typical examples are the brain, the immune system, the economy, social systems, ecology, insects swarm, etc.
>
> Therefore to understand a complex system one has to study the system as a whole and not to decompose it into its constituents. This totalistic approach is against the standard reductionist one, which tries to decompose any system to its constituents and hopes that by understanding the elements one can understand the whole system. (Ahmed et al. 2005, pp. 1–2)

According to Murray Gell-Mann (1992, 1994, 1995), the category of CAS should also include all the basic components of this system as well as individuals who are capable of surviving by adapting their behavior and producing new *schema of interaction* and coexistence that allow this behavior to be *predicted* and adapted to.

> Now how does a complex adaptive system operate? How does it engage in passive learning about its environment, in prediction of the future impacts of the environment, and in prediction of how the environment will react to its behavior?
>
> [...] The answer lies in the way the information about the environment is recorded. In complex adaptive systems, it is not merely listed in what computer scientists would call a look-up table. Instead, the regularities of the experience are encapsulated in highly compressed form as a *model* or *theory* or *schema*. Such a schema is usually approximate, sometimes wrong, but it may be adaptive if it can make useful predictions including interpolation and extrapolation and sometimes generalization to situations very different from those previously encountered. In the presence of new information from the environment, the compressed schema *unfolds* to give prediction or behavior or both. (Gell-Mann 1992, p. 10)

In addition to what Gell-Mann had to say, I would observe that CAS have been recently studied in the context of the theory of complex systems. Briefly, I would note that the CAS *approach* studies how complex systems interact and exchange information with their environment in order to maintain over time their internal structure and the network of vital processes and to develop a form of social cognition. The term CAS refers to a system (not exclusively of a social nature but of an organizational one as well) with the following properties:

– It is composed of a large number of primitive components, or "agents," different in nature (men, animals, plants, robots, scientific theories, neurons, etc.);
– Their number is not always fixed, so that the system can often be considered open; thus, it may be difficult or impossible to define system boundaries;
– It produces many types of different interactions among the agents and between the agents and their environment; these interactions are in the form of reinforcing and balancing loops, which make the interactions *nonlinear* since, as we know, small actions by the agents can produce significant changes in the system;
– The agents are structurally coupled to other agents and to the environment, and they are subject to many environmental constraints;
– As a consequence of the interactions among the agents, the system's behavior evolves over time;
– Unanticipated *global properties or patterns emerge* as a result of often nonlinear spatial–temporal interactions among a large number of component systems at different levels of organization.

There are three basic characteristics of a CAS which allow us to interpret and understand the capacities for self-control and lengthy survival in social systems:

(a) a CAS is comprised of *different types* of agents that try to adapt their behavior to that of the other agents they can observe adapting to one another "...*as a process of reciprocal selection of congruent paths of structural changes in the interacting systems which result in the continuous selection in them of congruent dynamics of state*" (Maturana and Guiloff 1980, p. 139);
(b) the agents' micro behavior is meant to adapt to disturbances only by following certain *local schema* (rules and strategies), not necessarily uniform across all agents, which can change over time when this favors adaptation; "*More generally, it is significant that any* CAS *is a pattern-recognition device that seeks to find regularities in experience and compress them into schemata*" (Gell-Mann 1994, p. 22).
(c) the processes of adaptation to environmental perturbations generates a succession of macro mutations in the system that enables it to *evolve* over time, thereby producing an evolutionary "history"; the dynamics of this evolution develop from the reciprocal local interactions according to local rules and strategies established over time or to variants produced by individual adaptation. Normally the agents are not aware of the macro behavior of the entire system, adapting instead to that of their neighbors within a restricted range of knowledge.

To break away from the paradigm of top-down modeling, we enter the science of complexity and its applications in complex adaptive systems theory. Instead of the top-down approach so entrenched in contemporary simulation and modeling, complex adaptive systems theory is based on the notion of model building from the bottom up. After three hundred years of dissecting everything into molecules and atoms and nuclei and quarks, they [scientists] finally seemed to be turning that process inside out. Instead of looking for the simplest pieces possible, they were starting to look at how those pieces go together into complex wholes (Waldrop 1992, p. 16).

Complex systems are usually systems which have been created by evolution or an evolutionary process. Evolved systems which have a long historical background are nearly always complex. Complexity can be found everywhere where evolution is at work,

- in all living organisms which are subject to evolution
- in all evolving complex adaptive systems which have a long history
- in systems that have grown over a long period of time. (CASG 2013, Online)

The term *complex evolving system* may be used to distinguish human from other CAS. In particular, *complex evolving systems* refers to those systems which are able to learn and which change their internal structure and organization over time, thus changing the behavior of the individual elements (Allen 1997). In social systems in particular, mutual adaptation among the individuals–agents not only is due to adaptation that is a reaction to the disturbances but derives from *forecasts*, *prospects*, and *expectations* regarding the behavior of the other individuals–agents, as is well stated by John Holland, the pioneer of CAS studies:

Here we confront directly the issues, and the questions, that distinguish CAS from other kinds of systems. One of the most obvious of these distinctions is the diversity of the agents that form CAS. Is this diversity the product of similar mechanisms in different CAS? Another distinction is more subtle, though equally pervasive and important. The interactions of agents in CAS are governed by anticipations engendered by learning and long-term adaptation. (Holland 1995, p. 93)

It is easy to understand that the process of *individual adaptation* presupposes the action of innumerable *Rings* operating recursively, at different levels, among pairs of individuals and/or groups to allow the agents to recognize the dynamics and construct the *schema* (models which are stable but adaptable to the circumstances) to achieve a state in conformity with them, as shown in Fig. 7.21. This model also enables us to see that the activity of each *complex adaptive agent* is conditioned by the agent's capacity to form and direct the evolution of the *schema* of adaptation on which the chosen coordination strategy depends, in order to adapt his behavior to the changing behavior of the other agents.

The repeated use of the innumerable *Rings* needed for adaptation and the continual modification of the *schema* of adaptation by all agents produces a form of *social learning*, which also makes increasingly more efficient and effective the action of control systems that operate recursively in order to allow the agents to achieve their individual objective through mutual adaptation in reaction to the observed behavior of the other individuals in the social system.

[…] social learning is based upon either extrinsic or intrinsic factors or motivations. Intrinsic social learning occurs in social facilitation. It occurs when people are led by others to acquire new means (intrinsic) for achieving old goals. Extrinsic social learning
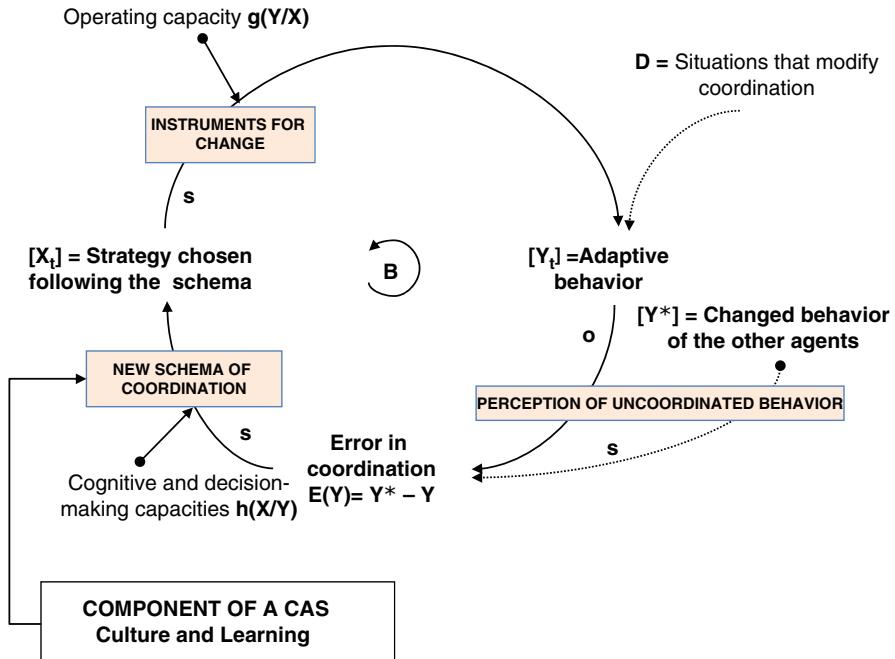
Fig. 7.21 The dynamic coordination of the agents in a CAS

occurs in imitation, when people learn through the observation of attractive and consistent social models. By observing their social models and recording when these apply reinforcing mechanisms (extrinsic), people learn to reinforce themselves (self-reinforcement) to do what others have reinforced, and abstain from doing what others have punished. (Conte and Paolucci 2001, par. 2.2)

CAS operate in every context and provide interpretative models of the numerous varieties of phenomena in the world. It is useful to propose the excellent example of CAS presented by John Holland in the following extensive quote:

On an ordinary day in New York City, Eleanor Petersson goes to her favorite specialty store to pick up a jar of pickled herring. She fully expects the herring to be there. Indeed, New Yorkers of all kinds consume vast stocks of food of all kinds, with hardly a worry about continued supply. This is not just some New Yorker persuasion; the inhabitants of Paris and Delhi and Shanghai and Tokyo expect the same. It's a sort of magic that everywhere is taken for granted. Yet these cities have no central planning commissions that solve the problems of purchasing and distributing supplies. Nor do they maintain large reserves to buffer fluctuations; their food would last less than a week or two if the daily arrivals were cut off. How do these cities avoid devastating swings between shortage and glut, year after year, decade after decade? The mystery deepens when we observe the kaleidoscopic nature of large cities. Buyers, sellers, administrations, streets, bridges, and buildings are always changing, so that a city's coherence is somehow imposed on a perpetual flux of people and structures. Like the standing wave in front of a rock in a fast-moving stream, a city is a pattern in time. No single constituent remains in place, but the city persists. To enlarge on the previous question: What enables cities to retain their coherence despite continual

disruptions and a lack of central planning? There are some standard answers to this question, but they really do not resolve the mystery. It is suggestive to say that Adam Smith's "invisible hand" or commerce, or custom, maintains the city's coherence, but we still are left asking How? (Holland 1995, p. 1)

Holland's answer is that, naturally, the phenomenon described is an example of order, which can be interpreted as the action of a CAS. From these few observations we can easily see the basic differences between combinatory systems and CAS.

Firstly, combinatory systems do not necessarily present phenomena of adaptation but, generally, some form of self-organization due to the *micro–macro feedback*, which is the adaptation of agents to a synthetic variable produced by the macro behavior of the system. A *second* difference is observable also as regards the similarity of the agents. Combinatory systems are made up of similar agents, while, as Holland notes, CAS are composed of heterogeneous agents: "*Here we confront directly the issues, and the questions, that distinguish* CAS *from other kinds of systems. One of the most obvious of these distinctions is the diversity of the agents that form* CAS" (Holland 1995, p. 93).

The *third* main difference regards the absence of interactions among the agents; in combinatory systems agents normally interact only with some macro variable and not each other. The *fourth* relevant difference is that the theory of CAS observes the macro effects of the system produced by the agents that follow a *schema* or change the schema previously followed. Any micro–macro feedback between the micro behaviors and the schema is considered as a relevant characteristic. *Finally*, ignoring the micro–macro feedback implies that CAS theory only focuses its attention on *necessitating* factors and ignores the *recombining* ones.

Taking account of these differences, we can conclude that combinatory systems can be viewed as a particular class of CAS. Nevertheless, combinatory systems can represent, precisely due to their simple logic, which can easily be observed in nature and easily described and simulated by using combinatory automata, a useful theoretical model to be applied in the study of populations of similar agents, though they are not applicable to complex social systems (Sect. 7.13.1).

7.12   Change Management in a Complex World: The PSC Model

Populations, societies and collectivities are continually evolving complex systems made up of interconnected social, political, and economic entities (holons) that form a network (holarchy) of social, political, and economic processes that produce a continuous exchange of goods, knowledge, and value in order to regenerate the autopoietic social system (Sect. 7.5). In this interconnected world, the basic components of society are ceaselessly regenerated and the network continually redesigned; this continuous process of adaptation produces an inexorable *change* in the entire society.
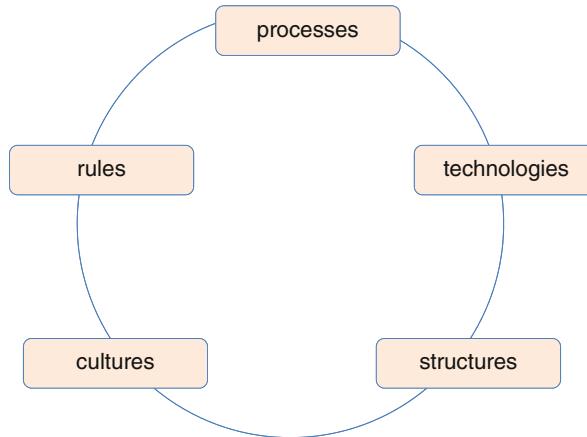
Fig. 7.22 The aspects of change in social environments

> Change in the substance of any of the relationships affects the overall structure. Since a change in any relationship affects the position of those involved, the whole set of interrelated relationships is subject to change and that has consequences for the outcome of a relationship for those involved. A dyad, a relationship, is a source as well as a recipient of change in the network. [...] (Håkansson and Snehota 1995, p. 41).

> We live in a changing world. People grow, business conditions fluctuate, personal relationships develop, major industries come and go, population increases, water shortages worsen, and scientific discoveries increase. Our principal concerns arise from change – growth, decay, and fluctuation. (Jay Wright Forrester 1983, Foreword, p. IX)

Traditional processes are abandoned while other processes are added, spurred on by the *new technologies*, which change the way social units and organizations operate (Hayes and Jaikumar 1988) according to the model in Fig. 7.22. Not only do the cultural bases of individuals, societies, and organizations change, but the social units also evolve through modification of their structures, processes, and output, in the attempt to loosen the old restraints while setting new objectives and rewriting the programs for their achievement.

Change to the culture of a social system and of an organization assumes an individual and social *learning* process that not only depends on the motivations for change but also influenced by the "living" and the "personality," as Talcott Parson clearly states:

> The crucial point for the present is that the "learning" and the "living" of a system of cultural patterns by the actors in a social system, cannot be understood without the analysis of motivation in relation to concrete situations, not only on the level of personality theory, but on the level of the mechanisms of the social system.
>   There is a certain element of logical symmetry in the relations of the social system to culture on the one hand and to personality on the other, but its implications must not be pressed too far. The deeper symmetry lies in the fact that both personalities and social systems are types of empirical action system in which both motivational and cultural elements or components are combined, and are thus in a sense parallel to each other. (Parsons 1951, p. 17)

Without going too far back into economic history, we can observe that social units and structures, organizations in particular, must deal with change in their environment. Business organizations, in particular, must operate in a more competitive context of heightened change and growing globalization (Porter 2008). Thus, in order to survive in a dynamic and changing world, social units and organizations must develop their ability to quickly understand change and react to unfavorable triggers, developing an intelligent behavior; that is, they must transform themselves into learning organizations (Senge 2006) by adapting or innovating their own structures and, above all, their own cultures. The change management process is conceived of as:

> The systematic approach and application of knowledge, tools and resources to deal with change. Change management means defining and adopting corporate strategies, structures, procedures and technologies to deal with changes in external conditions and the business environment. (SHRM Online)

This process becomes physiological, for individuals (Lewin 1947a, b), social groups, or organizations (Senge 2006), and it represents the natural approach for dealing with change both at the individual and organizational levels for all types of organization (Hiatt 2006).

For Kurt Lewin:

> A change towards a higher level of group performance is frequently short lived; after a 'shot in the arm,' group life soon returns to the previous level. This indicates that it does not suffice to define the objective of a planned change in group performance as the reaching of a different level. Permanency at the new level, or permanency for a desired period, should be included in the objective. (Lewin 1947a, p. 228)

For this reason Lewin proposed a three-step model of change:

Step 1: *Unfreezing*. Unfreezing from a position of harmful equilibrium that impedes change in groups is necessary to create the motivation for change; however, this does not necessarily produce and control change. Unfreezing is the necessary condition for overcoming the barriers to organizational change and initiating change (Argyris 1993).

Step 2: *Moving*. Even when the situation to change is unfrozen, the directions of change are not always clear. The control process toward new desired states appears difficult, and for this reason Lewin recommends that the groups move forward gradually, recognizing all the forces at work and identifying and evaluating, on a trial and basis, all the available options. This is the perfect description of a control system for change, which produces a gradual shift to move a group from a situation that has been unfrozen to a new, evolving one. Change management is the management of this control system, which must move toward the objectives of change established by the governance.

Step 3: *Refreezing*. In this third step the results of the changes judged to be satisfactory must be refrozen in order to stabilize the group at a new quasi-stationary equilibrium in order to maintain the new state relatively safe from regression.

The above considerations for group and organizations can refer, on a smaller scale, even to individuals. The survival of man, whether as an individual or as part of an organized group, depends on his ability to understand and dominate the changes in the variables that influence his choices and behavior, and he must develop the capability of being part of a network of knowledge and ideas. Therefore, such cognitive activities are very difficult today because change has become "complex"; individual, collective, and environmental dynamics interact, giving rise to reinforcing and balancing loops that form "complex" systems of variables that produce outcomes which are less and less understandable and predictable. In the face of environmental pressures, the capacity of the social systems to survive at length—that is, to maintain themselves as viable systems (Beer1979) and develop teleonomy based on Jacques Monod's (1970) conception—depends on the ability of the managers of change to understand the change and regenerate the internal vital processes. According to Ross Ashby's well-known requirement of *necessary variety* (Sect. 5.3) in a world of increasing dynamics and complexity all self-directed social systems must be "reviewed" and strengthened through an effective action of *change management*.

Change management is not a simple process but involves several fundamental phases that can be listed as follows:

1. Ascertaining the level of dissatisfaction with the present situation and the need for change;
2. Recognizing the optimal situation to achieve (vision);
3. Identifying the paths of change;
4. Structuring a strategy of change;
5. Identifying the person in charge of the process and setting up the task force of change management;
6. Undertaking change and developing the actions of change;
7. Controlling the adequacy of the implemented changes in order to achieve the desired optimal situation;
8. Stabilizing the changes obtained if these are held to be satisfactory, or returning to phase 2.

Apart from the general method of the three steps proposed by Kurt Lewin, different strategies have been presented to direct the fulfillment of these steps—the ADKAR model (Hiatt 2006), the McKinsy "7S" model (Waterman et al. 1980), Kotter's model (Kotter 1996) as well as others—but all models show that the manager of change should have the ability to recognize and meet the demands of change, possess the will to make change happen, and gain the consent of the parties involved in the change. Interpreted from the control system view, these models lead to the recognition that change management is a general *control process* to *move the system to be changed* toward the objectives that allow them to survive.

The phases for carrying out a change management process are by now standardized and well-consolidated in doctrine, but I feel it is useful to reconsider the central phases in the process, which are indicated in points 3 and 4, in order to place them in a dynamic pattern.

Cognitive models, attitudes and behavior change
social and educational structures are involved
long time periods, great difficulty

**Structures, instruments
and people change**
investment, difficulties low speed

Processes and procedures change
high-speed wheel
facility, low cost

New behavioral and
cognitive models of
the individuals or
groups

New **structures**
needed
for change

Rules which
oblige to
modify
processes
and routines

EASY REVERSIBILITY

LOW REVERSIBILITY

IRREVERSIBILITY

Fig. 7.23 The PSC model showing the wheels of change in social systems

With regard to identifying the paths of change of the *vital processes* of every social system, we must recognize there are three ways to realize change, each of which identifies a different lever that operates at different speeds:

1. the change in the internal operational *programs*, understood also as norms, regulations, instructions, etc.;
2. the change in the *structures*, that is, in the organization of the organs and individuals;
3. the change in the *culture*, that is, in the cognitive and behavioral models of the individuals or groups that participate in the social and organizational structure.

The three paths of change are interconnected, like wheels of an inexorable process of change as shown in the PSC Model (Processes, Structures, Culture) of Fig. 7.23; however, they lead to change at different speeds (Mella and Meo Colombo 2012).

Considering the rotation speed, the small wheel of *programs* that regulate the processes is the fastest but also the least effective. It spins easily and quickly. Nothing is easier than changing a norm, a regulation, placing a ban, and limiting the alternatives. Any program change can be modified by a subsequent change. The speed of the small wheel's rotation assures a timely adjustment to environmental dynamics. However, the small wheel causes the medium wheel to spin, even though slowly. This form of speedy change, at relatively low cost, is typically symptomatic and lends itself to continual revisions, since it is easily reversible and more affected by the temporary equilibriums among the participants in the governance.

The change of an operational *program* imposes a change in the *processes*, which is often traumatic for the structure that is thus forced to adapt, slowly, in order to respect the modified programs; the structure learns the new changes and itself changes. Through a new regulation we speedily introduce new subjects in scholastic programs, modify the structure of university degree programs, place new bans or limits on the circulation of people and cars, limit the sale of certain goods, impose/ban the use of additives, change the rules at work, forbid the turning on of boilers or air conditioners, forbid smoke in certain locales, etc. All these sudden changes produce a variation in the structures that must carry them out: schools must recruit the appropriate teachers, the universities must reorganize their faculties and departments, the traffic wardens must put up new signs, changes must be made in packaging and distribution processes, smoke ventilators must be installed and smoking corners set up, and so on. However, the changes in the structures that must adjust to the new processes proceed much more slowly than do the changes in the programs and processes themselves, since changes in the operational structure of the institutions and organizations are very costly; moreover, once carried out, structural changes are more difficult to change, becoming almost irreversible. The return to the *status quo* is nearly impossible, since the institutions and organizational system have been deeply modified (Watzlawick et al. 1974). Though slowly, the wheel of change in structures turns the largest wheel, one that of cultural change, and this slow change produces cultural innovation, which modifies the very foundations of society and of the organization. Once in place this process is nearly irreversible, since any change to it would require an equally long period for the change itself.

> Changing people's habits and way of thinking is like writing your instructions in the snow during a snowstorm. Every 20 minutes you must rewrite your instructions. Only with constant repetition will you create change. (Donald Dewar, cited by Haines 2005, p. 53)

This third level of change also includes the change in social and organizational values, which Casey (2000) has defined as the third order of change. Cultural change can be quite inconvenient for social systems, since many individuals could find it difficult to accept, adapt to and live the new cultural values. Thus, cultural change could produce a true evolution in the entire social system or in its large component parts, as well as a breaking up of the system into subsystems that adopt the old and new cultural and value structure. Cultural change represents the definitive solution that exploits the potential of informed and educated individuals. Without such change the normative efforts (small wheel) and the structural ones (medium-size wheel) run the risk of being merely symptomatic interventions on individual variables, and thus of not producing the *leverage effect* needed to deal with sustainability.

We arrive at these speculative conclusions by also considering the slowest wheel first. When a society or an organization succeeds in effectively influencing the culture of those belonging to it, then the new cognitive models (large and slow wheel) trigger a fast change in the structure (middle wheel), which, in turn, produces an ever faster adaptation in the programs and procedures (smallest wheel).
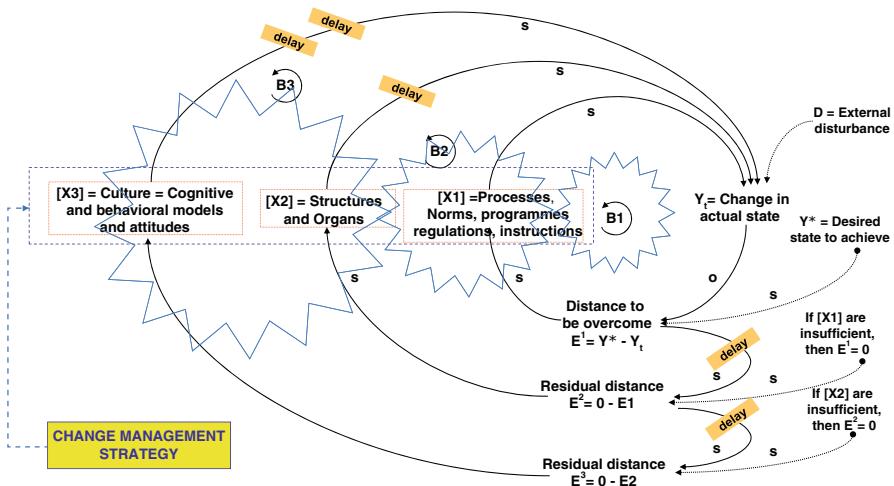
Fig. 7.24 The control of change through a multi-layer *Ring* representing the PSC model

In Fig. 7.24 it is easy to recognize the three wheels of change of Fig. 7.23, a typical system of multi-layer control.

The control process is realized through a *problem finding* and *problem solving* process (Nickols 2010, 2011). Considering change management as a particular form of creative problem solving, the two views, which we have called *top down* and *bottom up*, can, however, also be considered as true *strategies* of creative change (Osborn 1953):

– the *symptomatic strategy*, which proceeds top down, from fast to slow, contingent to permanent;
– the *permanent strategy*, which proceeds bottom up, from slow to fast, permanent to contingent.

Environmental changes, which are accelerated and interconnected, often affecting symbolic variables, produce unpleasant effects which the social systems interpret as *problems*, as *symptoms* to eliminate as quickly as possible. The symptomatic strategy (top down), which is reactive and quickly carried out, is considered an immediate, though contingent, remedy for dealing with unpleasant problems due to change. Change management must act rapidly, making the most urgent changes to the present situation, often without involvement of others in the organization, by acting on rapid levers such as programs, directives, and operational rules.

As the model in Fig. 7.25 shows, the symptomatic solutions, though fast, almost always lead to collateral effects that can worsen the problems arising from change. The *permanent* strategy (bottom up) is the most extensive, and for this reason it is slow and difficult to achieve, since it forces change on the cultural wheel by spinning the wheel of cultural change; but if it is successfully adopted, then the social system will have reacted to the change in a lasting way and have become able to deal with other rapid changes.
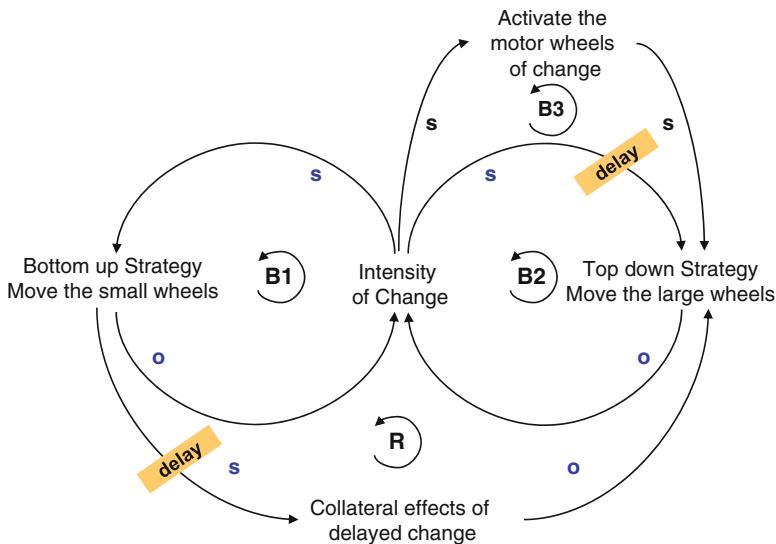
Fig. 7.25 The symptomatic and permanent solutions of change management

To quote Peter Senge:

> The advantage of systems thinking derives from the leverage effect – seeing in what way the actions and changes in the structures can lead to long-lasting, meaningful improvements. Often the leverage effect follows the principle of the economy of means, according to which the best results do not come from large-scale efforts but from well-concentrated small actions. Our non-system way of thinking causes significant specific damage because it continually leads us to concentrate on low leverage effect changes: we concentrate on symptoms of higher stress. We correct and improve the symptoms: but such efforts are limited, when things go well, to improving short-term factors, while worsening the situation in the long run. (Senge 2006, p. 131)

In this way not only can the system forbid alcohol and counter the spread of new drugs in order to reduce "Saturday night road deaths," but change the smoking and drug culture through a long-term strategy of education and the self-fulfillment of youth; not only limit, by law, toxic emissions or change the structure of productive systems but introduce an appropriate cultural attitude to environmental problems.

> Change is a central concern in every human activity… But the processes of change have not been presented in an orderly way in our educational institutions. The dynamics of change have seldom been taught as a basic foundation that underlies all fields. The processes of change have not been organized so that they can be taught at all educational levels, even though a child, from his or her earliest awareness, begins to cope with change and to build an intuitive awareness of change. (Jay Wright Forrester 1983, Foreword, p. IX)

Of the three paths of change in social systems, that which leads to a cultural adaptation appears, though slow to occur, the most necessary today, precisely due to its propulsive power regarding the other wheels of change in a world that is changing more and more rapidly and in a structurally complex manner.

7.13   Complementary Material

7.13.1   *Models and Classes of Combinatory Systems*

The typical logic of the combinatory systems outlined in Fig. 7.17 produces a wide class of collective phenomena; that logic can be generalized by introducing these simple control rules that direct the micro and macro behaviors:

1. each agent is characterized by an *individual variable* of some kind (qualitative or quantitative) whose values—at any time t_h—represent the *individual states*;
2. the collectivity is characterized by a *global variable* (qualitative or quantitative) whose values—at any time t_h—represent the *system state*;
3. due to the presence of an opportune set of *recombining factors*, the system state—at any time t_h—derives from the *combination* (to be specified) of the individual states, following macro or *recombining rules*;
4. each agent can perceive a gap (positive or negative) between his individual state and the state of the collectivity;
5. due to the presence of an opportune set of *necessitating factors* each agent—at time t_{h+1}—decides, or is forced, to identify the most appropriate strategies to expand or reduce the perceived gap following the micro or *necessitating rules*;
6. as long as the necessitating and recombining factors are maintained, the *micro–macro feedback* can operate;
7. the agents are characterized by an initial state at time t_0; in most cases this initial state may be assumed to be due to *chance*;
8. the micro–macro feedback operates between the limits of the *minimum activation number* and the *maximum saturation number* of the agents who reveal the state maintaining the micro–macro feedback;
9. the sequences of state values over a period represent the macro and micro dynamics, or behaviors, of the system and the agents, as shown below in Fig. 7.26.

Combinatory systems can be represented by different models of increasing complexity. The simplest models are the *descriptive* ones that indicate in words—or by patterns analogous to Fig. 7.17 (energy inputs are not included)—the fundamental elements necessary for understanding the operative logic of systems that produce observable collective phenomena. The more powerful models are the *heuristic models* that try to simulate the system's dynamics by stating—or constructing ad hoc—a *set of rules* specifying: (a) the micro, or *necessitating* rules producing the micro behaviors of agents as the consequence of the necessitating factors; (b) the macro, or *recombining* rules that produce the system's macro behavior, due to the presence of recombining factors; (c) the *micro–macro feedback* that allows the system to produce the observed phenomena; (d) the *strengthening*, *weakening* and *control levers*, when possible or admitted. Finally, we can build a combinatory automaton (Mella 2007) that specifies the mathematical and statistical simulation model that produces the micro and macro behaviors of the combinatory systems based on specific assumptions regarding the *recombining* and *necessitating* rules.
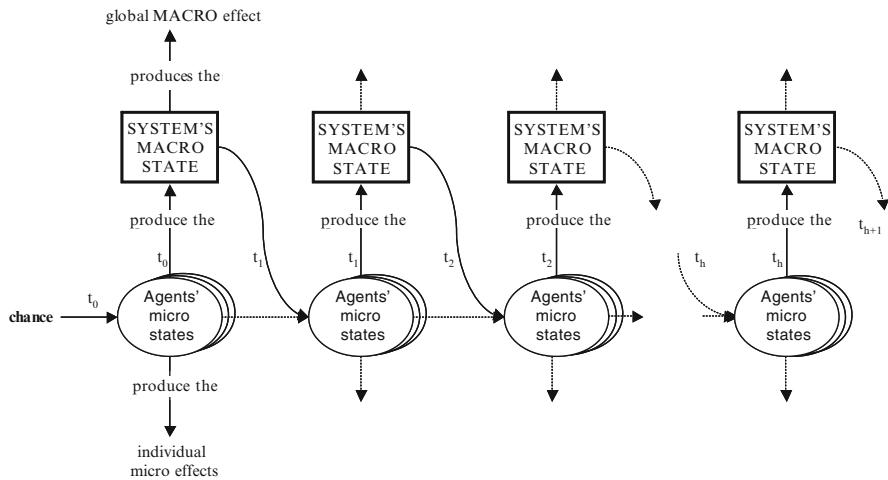
Fig. 7.26 The micro and macro dynamics and the micro–macro feedback

The *heuristic models* of the five important classes of combinatory systems (described in Sect. 7.8) can be briefly described as follows; a comparison between them clearly reveals the same *modus operandi* of the combinatory systems of the various heuristic models.

1. Heuristic model for systems of *accumulation*.

    NECESSITATING RULE: if you have to accumulate some object with others similar in nature (micro behavior), look for already-made accumulations, since this gives you an advantage or reduces some disadvantage (necessitating factor).

    RECOMBINING RULE: the environment preserves the accumulated objects, or is not able to eliminate them, thus maintaining the advantages of the accumulation; everyone accumulates (macro behavior), and an accumulation of some kind is created (macro effect).

    MICRO–MACRO FEEDBACK: the accumulation is the macro effect of single micro behaviors; the larger the accumulation (macro effect), the more incentive (facility, probability) for accumulating (micro behaviors) objects (micro effects); the collective accumulation (macro behavior) leads to the mainte-nance or the increase of the accumulation.

    STRENGTHENING, WEAKENING AND CONTROL LEVERS: a sign reading "accumulate here" or "accumulation prohibited" represents the best strengthening action. The prompt removal of the initial accumulation and a careful vigilance represent weakening factors. The external control acts on the accumulation and on the site where this forms; the internal control acts on the desire to accumulate *and* on the individuals' actions.

2. Heuristic model for systems of *diffusion*.

   NECESSITATING RULE: if you see that an "object" is diffused, then it is "useful" for you to possess it or harmful not to possess it (necessitating factor), and you must try to produce or acquire it.

   RECOMBINING RULE: the environment or the collectivity preserves the diffused objects and maintains the utility of possessing the object; the higher the individual's utility or need is to acquire the object, the more the object will spread throughout the collectivity.

   MICRO–MACRO FEEDBACK: a greater diffusion (macro effect) implies a greater desire to acquire the object (micro effect); the single acquisition (micro behavior) increases the collective diffusion (macro behavior).

   STRENGTHENING, WEAKENING, AND CONTROL LEVERS: publicity and social gratification represent strengthening factors for the system; social disapproval and repression represent weakening factors; the macro control must act to limit the availability of purchasable goods; the micro control must act on individual preferences.

3. Heuristic model for systems of *pursuit*.

   NECESSITATING RULE: if there is an objective, try to achieve it; if there is a limit, try to exceed it; if another individual overtakes you (negative gap), regain the lost ground; if you are even with someone, try to go ahead of him; if you are in the lead, try to maintain or increase your advantage (positive gap).

   RECOMBINING RULE: the collectivity recognizes the validity of the objective and views limits in a negative way; the more individuals try to exceed the limit, the greater the chance of exceeding it, with a consequent advantage for those who succeed in doing so. This provides the incentive for the pursuit.

   MICRO–MACRO FEEDBACK: if everyone tries to go beyond the limit (macro behavior), then this limit is raised (macro effect), thereby eliminating the advantage for those who have already reached it (micro effect); this forces individuals to exceed the limit (micro behavior).

   STRENGTHENING, WEAKENING, AND CONTROL FACTORS: if honor and prize money are awarded, then the system is strengthened due to the increase in gratification. If the achievement of the objective or the exceeding of the limit causes accidents, with victims, then there is a weakening. Another weakening factor involves the cost of competitions: the higher this is, the more difficult the search for the record becomes. Control factors for the system are represented, for example, by publicity concerning the record and the creation of venues where the sporting activity can be practised. A typical micro control intervention is represented by a favorable culture for competitions.

4. Heuristic model for systems of *order*.

   NECESSITATING RULE: there are advantages in maintaining a particular order and disadvantages in breaking it; if you want to gain advantages or avoid disadvantages, try to control your behavior so that you maintain or achieve the order that is indicated by the rules that establish it.

RECOMBINING RULE: the more the particular order is maintained, the greater the advantages from adjusting one's behavior to maintain it and the disadvantages from breaking it.

MICRO–MACRO FEEDBACK: the order (macro effect) creates the convenience for individuals to maintain the arrangement and respect the rules (micro behaviors); everyone maintains a coordinated behavior (macro behavior).

STRENGTHENINGS, WEAKENINGS, AND CONTROL: strengthenings to the system can be represented by awards for those who maintain the order and evident punishments for those breaking it. Introducing difficulties and constraints to the maintenance of order discourages individuals from respecting such maintenance. Facilitating and removing constraints make it easier for agents to control their micro behavior. The macro control can be carried out through the supervision of a director who forces the dance pairs to dance in a rotating way. The prior instruction represents a form of micro control.

5. Heuristic model for systems of improvement and progress.

NECESSITATING RULE: if you perceive that the level of your improvement parameter is below the level of the system's progress parameter—that is, there is a negative gap between your state and that of the others—try to improve in order to reduce the gap and, if possible, to attain a positive gap; if you perceive there is a positive gap, do nothing or try to improve further in order to increase the favorable gap.

RECOMBINING RULE: the system must be able to notice the individual improvement and adjust the progress parameter to the average (or, more generally, to the combination) of the individual improvement measures.

MICRO–MACRO FEEDBACK: individual improvement (micro effect) raises the parameter that measures collective progress (macro effect); this leads to the formation of positive and negative gaps that push the individuals to control their behavior in order to increase the gaps (if positive) or eliminate them (if negative).

STRENGTHENINGS, WEAKENINGS, AND CONTROL: the macro control can be carried out through a supervisor or a manager who makes the agents aware of their inferior state and of the possibility of improving their position, or advises them on how to do so. The prior instruction of the dance pairs represents a form of micro control.

7.13.2   The Combinatory Automaton for Simulating Buzzing in an Indoor Locale

In order to simulate the behavior of the different classes of *combinatory systems* and their macro effects, it is useful to build a simple linear *combinatory automaton*, made up of a grid (or an array) of cells, each of which contains the value of the micro state of an agent of the combinatory system. At each instant, t_h, the automaton calculates the synthetic state of the entire system through an operation (addition, average, max value, min value, etc.) on the values of all the cells. At the

Parameters of the model	symbols	A_1	A_2	A_3	...	A_{18}	A_{19}	A_{20}
Initial voice level of A_i	$v_i(t_0)$	1,00	1,00	1,00	...	0,00	0,00	0,00
Amplification coefficient	k =	0,8		a = 0	...			
External noise	Q =	10			...			
Position coefficient of A_i	w(n)	1,00	1,00	1,00	...	1,00	1,00	1,00
Decibels above R	$v_{i(min)}$	10	9	8	...	9	8	7
Decibels above R (random)	$v_{i(rnd)}$	2	3	4	...	3	4	1
Necessitating factor (probab.)	$s_{i\,[0,1]}$	60%	65%	70%	...	70%	80%	90%

Fig. 7.27 Control panel for the combinatory automaton that produces background noise

subsequent instant, t_{h+1}, the individual cells are updated by calculating a new state that takes into account the overall system state, thus producing the dynamics shown in Fig. 7.26. In combinatory automata the *recombining* and *necessitating* factors are translated into numerical probabilities that are assigned to the various cells and updated for each iteration by applying some rule.

The combinatory automaton model of the combinatory system that produces buzzing noise in a crowded room is shown in Fig. 7.27, from which we can also derive the logic of the simulation model that describes the macro behavior of the system (production of noise) and the micro behaviors of the agents (speaking at a voice level that exceeds the noise).

Taking into account these assumptions, I can represent the combinatory system by the following simple probabilistic model:

Input of the initial state of the system : $v_i(t_0) \leftarrow$ "chance", $1 \leq i \leq N$

$$R(t_h)] = \left\{ k \left[(1/N) \sum_{1 \leq i \leq N} v_i(t_h) \right] + Q\, r(t_h)_{[0,1]} \right\} (1 - a), h = 0, 1, 2, \ldots$$

$$v_i(t_{h+1}) = \left\{ \left[w_i R(t_h) + v_{i(min)} \right] + v_{i(rnd)}\, l_i(t_h)_{[0,1]} \right\} s_{i[0,1]}\, b_{i[bol]}(t_h).$$

The descriptive model of the combinatory system producing noise in crowded locales is shown in Fig. 7.28. The basic variables and parameters that characterize this combinatory automaton can be summarized as follows:

1. the group of speakers consists of N agents, A_i, with $1 \leq i \leq N$; their behavior is observed over discrete instants: t_h, $h = 0, 1, 2 \ldots$;
2. $R(t_h) = k \left[(1/N) \sum_{1 \leq i \leq N} V_i(t_h) \right]$ represents the background noise, the Buzzing that arises from the combination of the speakers' voice levels; we assume this is determined by the arithmetic average of the speakers' voices, which represents the principal *recombining rule*; the average is rectified by the amplifying factor "k," whose value depends on the shape of the room;
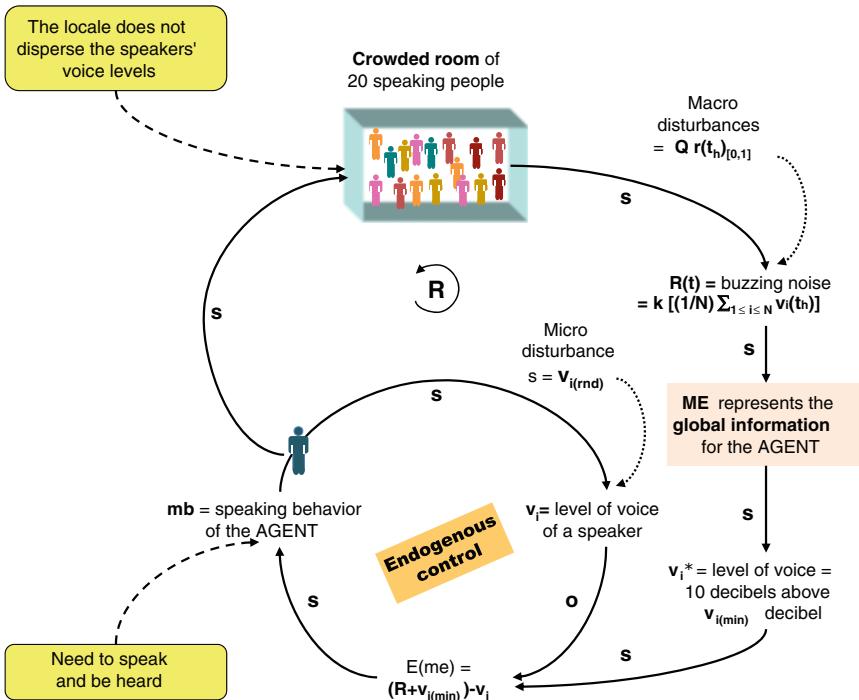
Fig. 7.28 Descriptive model for the combinatory system that produces background noise

3. the background noise, $R(t_h)$, does not depend solely on the speakers' voice levels but on other factors as well, among which the outside noise, which represents a macro-level disturbance; this is indicated by $Qr(t_h)_{[0,1]}$, with Q being the disturbance and $r(t_h)_{[0,1]}$ the probability of this occurring at each instantly have also considered the possibility the background noise is externally controlled by sound absorbing panels by introducing a coefficient of sound-absorption, "a"; if $a = 1$ there is total sound absorption and $R(t_h) = 0$;

4. $v_{i(min)}$ indicates the number of decibels (which differs for each agent) each speaker must produce above $R(t_h)$ in order to be heard by the others;

5. $v_i(t_{h+1}) = [w_i R(t_h) + v_{i(min)}] + v_{i(rnd)}$ represents the voice level of the agents above the buzzing level; the parameter w_i (usually set equal to 1) represents a synthetic measure of the position of each agent: if $w_i > 1$ the Buzzing is perceived as amplified by the speaker, who then further raises his voice;

6. for each agent an anomalous variation of the voice level is added, indicated by v_i (rnd), to take into account random factors that oblige speakers to modify their way of speaking (emotional state, distance from other speakers, etc.);

7. $s_{i[0,1]}$ indicates the probability of speaking; this is a very important variable in the model, since it synthetically expresses the complex of necessitating factors that push individuals to speak or remain silent; for simplicity's sake, this probability is assumed to be constant over time, even though it is different for each speaker;

8. I have also introduced tolerance to noise, indicated by $b_{i(\text{bol})}(t_h)$; this is a Boolean variable that is equal to "0" when the speaker has reached the maximum level of toleration and stops talking.

The preceding rules have been translated into a combinatory automaton of $N = 20$ agents talking in a room, each of whom is characterized by individual parameters and probabilities of speaking specific to each agent, which are represented in the control panel in Fig. 7.27.

The simulation is presented in Fig. 7.29, which shows the dynamics in the voice levels (colored lines) and the buzzing noise (bold blue line) for a period of 30 cycles of the process (each cycle hypothetically lasts 2 s).

The *internal control* is provided by the individual speakers, who vary their voice levels when the buzzing varies; typical forms of external macro control are the placing of sound absorbing panels on the walls to reduce the action of the recombining factor that transforms the voices into a buzzing noise, thereby modifying the coefficient "k," or keeping macro-level external disturbances under control.

A form of *micro control* is the education and upbringing of the speakers, which affect the probability of speaking, which we have assumed includes all the necessitating factors. Another form is to assign a supervisor to the room (room director, usher) who, at regular intervals, invites those present to moderate their voice levels.

During a performance in a concert hall only rude people whisper, and they are immediately told to be quiet by those nearby or are even reproached by the usher. The dynamics produced by the combinatory automaton in Fig. 7.28, which simulates different situations, is shown in the upper model in Fig. 7.29. The noise (bold line) may be viewed as the output of the combinatory automaton constituting the collectivity considered as a whole.

7.13.3   Two Modern Tragedies of the Commons

The ARCHETYPE of the "Tragedy of the Common Resources" (Sect. 7.10 and Fig. 7.19), which derives from the incessant and widespread action of the ARCHETYPE of "short-term, local and individual preference" (Sect. 1.6.6 and Fig. 1.19), produces worrisome appropriation phenomena which weakens the capacity of social systems and organizations to develop and even survive over time.

This has given rise to and increased a "social alarm" that sees in the micro behavior of agents guided by "short-term, local and individual preference" the main source of the problem concerning the *sustainability* of the macro behavior of social, economic, and environmental systems.

> Biologically, sustainability means avoiding extinction and living to survive and reproduce. Economically, it means avoiding major disruptions and collapses, hedging against instabilities and discontinuities. Sustainability, at its base, always concerns temporality, and in particular, longevity. (Costanza and Patten 1995, p. 196)

The problem of sustainability is usually considered looking forward and interpreted as a problem of the *sustainable development* of the macro variables of collective systems. In social systems, sustainability is usually defined as:

Test 1 – Classroom External noise Q =10 decibels. Mean probability to speak= 85%

Test 2 – Classroom External noise Q = 0 decibels. Mean probability to speak= 25%

Test 3 – Theatre during a performance. Q = 10 decibels. Mean probability to speak= 3%
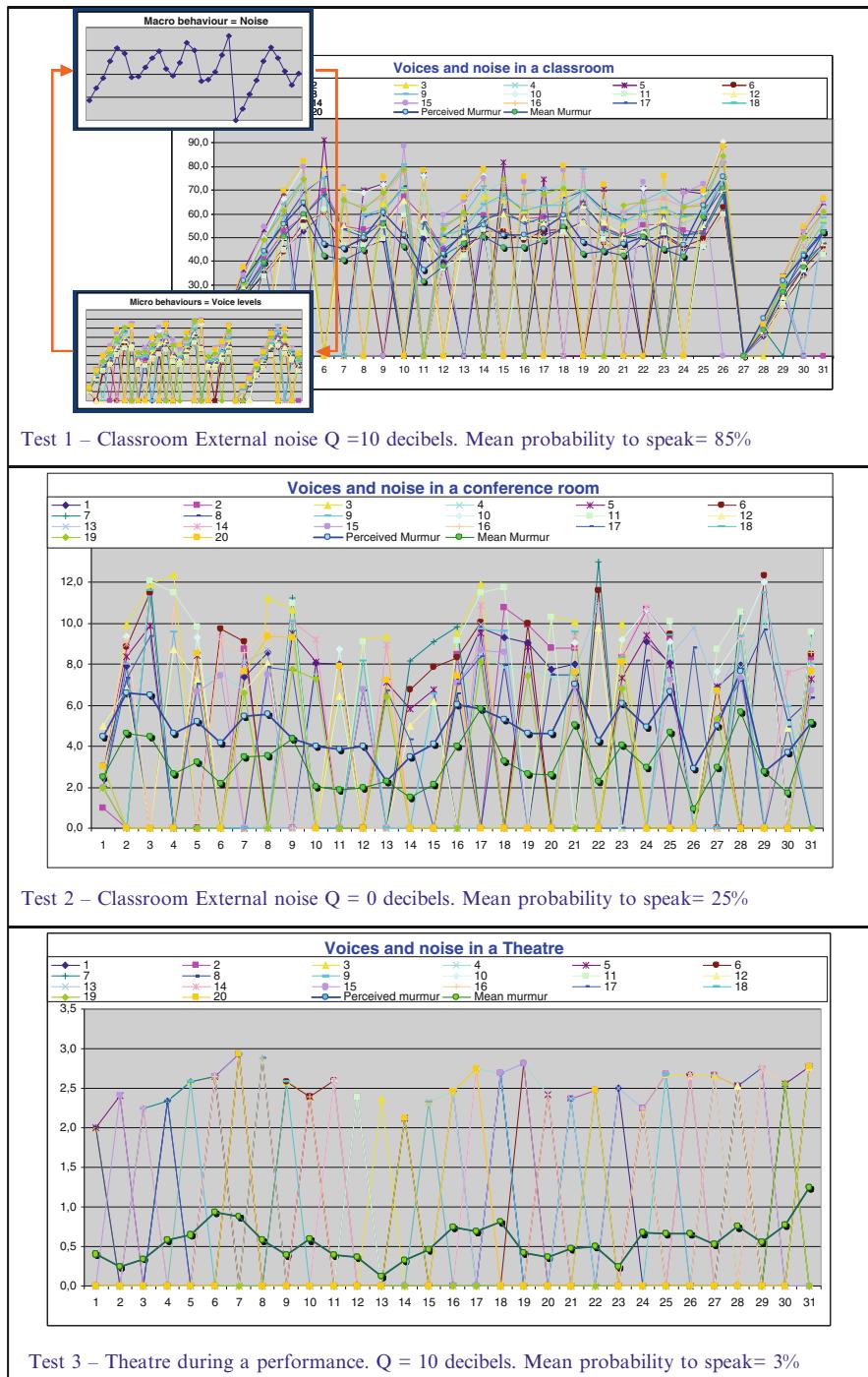
Fig. 7.29 Combinatory automaton simulating noise in different locales and situations (Powersim)

> [...] a broad interpretation of ecological economics where environmental and ecological variables and issues are basic but part of a multidimensional perspective. Social, cultural, health-related and monetary/financial aspects have to be integrated into the analysis. (Söderbaum 2008)

At the start of the 1970s an important report entitled "The limits to growth" (Club of Rome 1972) initiated the international debate on how man should intervene to create a curve of logistic accommodation to resources in order to limit the exponential growth in population, food, industrial production, energy consumption, CO_2 emissions, etc., and thus avoid a catastrophe. The nature of the debate shifted from "limits to growth" to the concept of "sustainable development" with the publication of the report: "World Conservation Strategy" (1980) by the International Union for the Conservation of Nature and Natural Resources (IUCN). Referring to the definition by the "Brundtland Commission" (1987), it has been observed that:

> Over these decades, the definition of sustainable development evolved. ...This definition was vague, but it cleverly captured two fundamental issues, the problem of the environmental degradation that so commonly accompanies economic growth, and yet the need for such growth to alleviate poverty. (Adams 2006, p. 1)

The most relevant and quoted definition of sustainability is the one from the report entitled, "Our Common Future," published by the WCED (institutional version):

> Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs. (Brundtland Commission 1987)

Thus, today "*there is wide consensus that the idea of sustainability figures as one of the leading models for societal development by indicating the direction in which societies ought to develop*" (Christen and Schmidt 2012, p. 411).

The archetype of "short-term, local and individual preference" (Fig. 1.19) reveals that the main problems of sustainability are caused by economic activity: individuals, as consumers, behave individualistically and with a short-term perspective due to their desire to consume goods in ever greater quantities and of ever higher quality; they do not know about, or are indifferent to, the long-run effects of their consumption. In fact, consumption requires heightened production activity, which, thanks to profit-motivated investment, which provides employment, earnings and taxes to finance public services, makes immediately clear the short-term and almost always local individual benefits. This repeated and increasingly intense production activity, and economic activity in general, produces three evident problems regarding sustainability:

(a) the depletion of nonrenewable resources;
(b) the depletion of disposal sites for waste and scraps of products that can no longer be used;
(c) increasing pollution, in particular global warming.
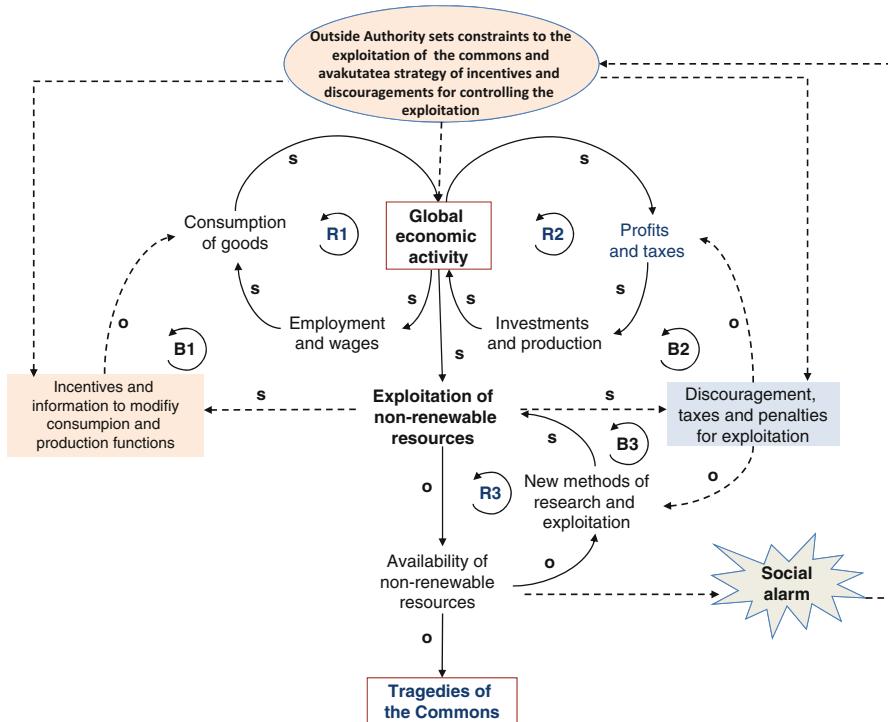
Fig. 7.30 Actions to counter the non-sustainability of the appropriation of the Commons and to avoid the Tragedies of the Commons

Point (c) has been discussed in Sect. 6.6. In this section I shall consider the first two problems. I begin with the observation that economic activity, which in the short term and due to corporate, national, or regional (local) interests, has a growing need for nonrenewable resources, becomes globally unsustainable in the long run and with a global dimension, because it depletes the nonrenewable resources.

The problem of the depletion of nonrenewable resources arises when a large number of firms, referred to as "appropriators" (indicated concisely in the model in Fig. 7.30 by the term "Global Economic Activity") and located in different areas, use a Common-Pool Resource that is available in limited quantities (the "commons") which are not distributed among the appropriators based on some indicator of merit or need but freely exploited in competition for the individual activities of the firms, whose outcome (production of goods, profits, wages, etc.) depends on the use of the resource. The loops [**R1**] and [**R2**] in the model in Fig. 7.30 show how the economic activity of the appropriators produces advantages for both consumers and producers-investors; as long as these advantages persist, economic activity will continue to increase.

It is obvious that the more the resource is exploited individually, the less is available collectively, so that even the activities that employ the resource are less

efficient, thus making its use increasingly "less economic." To achieve their final objectives in a situation of lesser efficiency the appropriators must intensify their activity, continuing to exploit the resource, no matter how scarce, until it is completely exhausted. The result is a "tragedy," since those who exploit the resource are forced to interrupt or reduce their activity because of its exhaustion. The efforts by the appropriators to pursue their objectives, which are individual and local and refer to a short-term time horizon, gives rise to more intense exploitation of the common-pool resources, as shown in loop [**R3**], whose action inevitably leads to the exhaustion of the common resource with negative effects for individual and collective outcomes.

If we observe the world around us we see this archetype in operation in many situations: too many fishermen fishing in waters with limited populations of fish; the search for oil, minerals and precious stones in the same area; and, in general, the exploitation of limited natural resources by many users. In conclusion, when the individual agents—consumers and producers—that produce the global economic activity need to exploit nonrenewable resources, without considering the sustainability of their activities, both with regard to consumers (eating whale meat may lead to the extinction of cetaceans) and producers (whaling vessels are now guided by satellite systems and can fire harpoons even from a distance of many kilometers in order to hunt down the few remaining whales), there is the inevitable risk such resources will be exhausted.

The action of the three archetypes summarized in Fig. 1.19 is inexorable; thus, educating producers and consumers regarding the question of long-term global sustainability has become increasingly more urgent, since the exploitation of limited common resources can cause conflicts that represent a "second tragedy." The clearest evidence of this is the numerous wars over water and oil, which are among the most limited and scarce vital goods on earth. How can the tragedies of the commons be countered? There is only one answer: when resources become very scarce a "social alarm" is sounded, and the public regulators, who become aware of the non-sustainability of the exploitation of limited resources, must externally control the behavior of consumers and entrepreneurs producing the harmful effects that must be avoided by specifying several useful actions to manage scarcity in order to reduce conflicts; some of these actions have been indicated in the model in Fig. 7.30 (see Fig. 1.20):

(a) Weaken loop [**R1**] by activating loop [**B1**], thereby encouraging consumers to change their tastes, thus directing their preferences toward products that do not use up scarce resources. This action on consumers would also affect producers, who would find it even more necessary and urgent to restructure their production functions by substituting scarce resources with non-scarce ones;

(b) Weaken loop [**R2**] by activating loop [**B2**], thereby discouraging producers by creating disadvantages for those who take advantage of the exploitation of nonrenewable resources. This action would force producers to pass on the greater costs to the prices of their products, which would discourage consumers from purchasing products using a large amount of scarce resources. The producers

should modify their production function and substitute the scarce resources with non-scarce resources;

(c) Weaken loop [**R3**] by activating loop [**B3**], thereby introducing disincentives to discourage innovation and proposing a moratorium on further exploitation of nonrenewable resources.

Other actions have been proposed alongside these three general ones, among which measures which directly impact the appropriators.

A second example that demonstrates the need to assess the sustainability of global economic growth concerns the problem of the disposal of waste, packaging material and products that are no longer usable (cars, appliances, electronic goods, chemical materials, etc.). Residential centers could not exist without the control systems for liquid and solid wastes, of which even the Romans were aware: even before any urban expansion occurred, they began enlarging their sewer systems. Today the production and distribution system for the necessity goods in life is also a powerful producer of solid wastes of all kinds that must be eliminated daily.

The general problem can be represented in the model shown in Fig. 7.31. Economic activity inevitably produces waste and by-products that, together with human biological waste, pose a serious threat to sustainability for future generations, since waste cannot always be disposed of and, even when it can, this implies (especially for solid waste) the use of increasingly larger areas for landfills and waste treatment sites.

The sustainability problem derives from the relative scarcity of such sites, which in turn results from the action of the archetype of local and individual preferences: communities with potential waste treatment sites refuse to accept the waste produced by inhabitants from other areas.

The areas housing landfills are in all respects Common-Pool Resources, and their gradual local exhaustion can give rise to the Tragedies of the Commons, exactly as described in the preceding section. Environmental disasters in the cities and geographical areas where the waste treatment sites are exhausted are ever more frequent, as well as the "wars" among inhabitants from differing areas who refuse to accept others' waste.

> Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons. Freedom in a commons brings ruin to all. [. . .] As a result of discussions carried out during the past decade I now suggest a better wording of the central idea: Under conditions of overpopulation, freedom in an "unmanaged commons" brings ruin to all. (Hardin 1985, Online)

Global regulators consider the scarcity of waste treatment sites as a problem that gives rise to a social alarm. In order to avoid the sustainability problem arising from the waste produced by global economic activity, global regulators can carry out four types of intervention. They can:

(a) concentrate on the local problem of waste treatment and try to convince the inhabitants of other geographical areas to accept waste produced elsewhere;
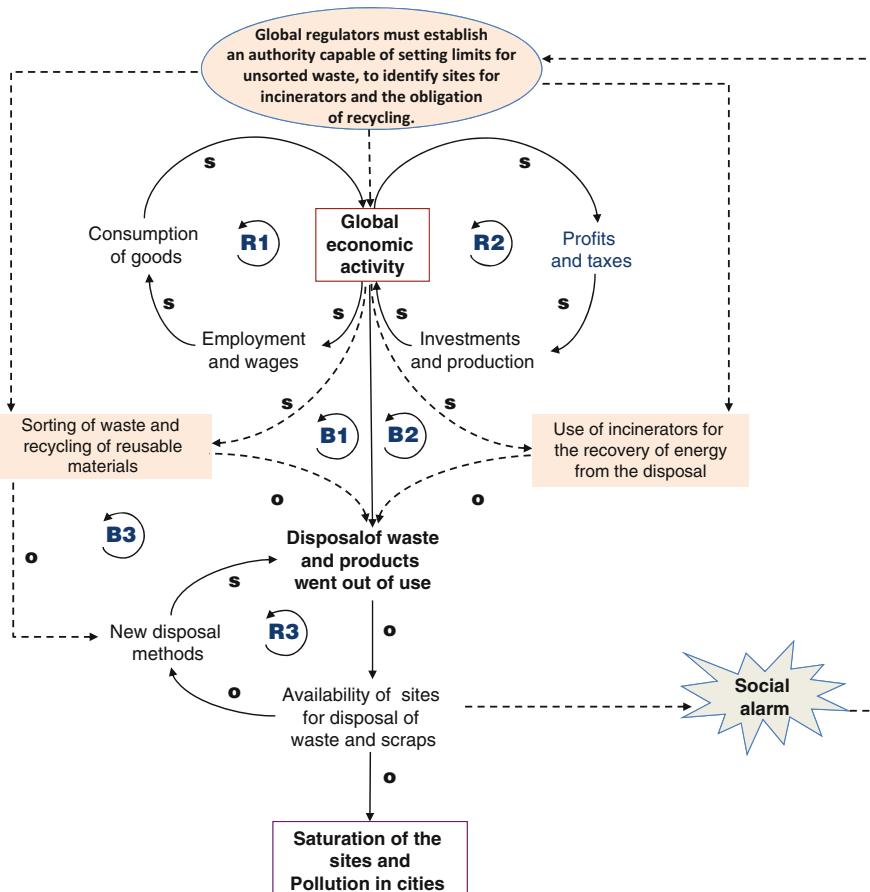
Fig. 7.31 The sustainability of pollution. Actions to counter the exhaustion of sites for the treatment of solid waste

(b) intervene in the waste production process through actions to discourage its production, taxes and penalties for polluting, and economic incentives for consumers and firms to induce them to modify consumption and production functions;

(c) provide incentives for the reutilization of waste or oblige individuals and firms to separate their waste collection and recycle their waste, thereby activating loops [**B1**] and [**B3**];

(d) modify the process of eliminating waste by providing incentives to, or obliging, individuals and companies to utilize waste in energy production and for heating purposes through incineration in increasingly more sophisticated, high-energy-efficiency incinerators; in this way loop [**B2**] is activated along, once again, with [**B3**] (the arrow is not indicated in the figure).

Figure 7.31 illustrates the functioning of measures (c) and (d), since the others operate exactly the same as those in the example in the preceding section.

Unfortunately, once again the archetype of local behavior can come into play, since the inhabitants of the areas where the sites would be located must be convinced to overcome their tendency toward localism and agree to separate waste collection and construct incinerators. Nevertheless, "*Not now and not in my backyard*" is still the most common response. Moreover, it is no surprise that orders from the regional authorities to limit traffic in a certain region are never accepted by all the cities in the region and that, for certain reasons, some local authorities permit a continual increase in the pollution thresholds.

7.14 Summary

Continuing our ideal journey among the *Rings* that make our world possible, this chapter has examined the control systems that characterize the macro biological environment, in particular:

(a) the *Rings* regulating the dynamics of interacting populations (Sect. 7.2), in particular the simulation of populations based on the Volterra–Lotka equations (Fig. 7.3), also including quantitative limits on the populations (Fig. 7.4);

(b) the *Rings* regulating the quantitative dynamics in ecosystems (Sect. 7.3) made up of multiple populations which form a food chain (Figs. 7.5 and 7.6);

(c) the *Rings*—both natural and artificial—regulating the qualitative dynamics of the populations (Sect. 7.4 and Fig. 7.8);

(d) the *Rings* that regulate the coexistence of individuals in social systems (Sect. 7.5 and Fig. 7.10);

(e) the *Rings* for the control of autopoietic variables (Sects. 7.6 and 7.7) based on the governance policy (Figs. 7.13 and 7.15);

(f) the *Rings* that allow the combinatory systems which form in collectivities to be activated (Sect. 7.8) and controlled (Sect. 7.9);

(g) the Tragedy of the Commons Resources, which occurs when the combinatory systems are composed of appropriators of a freely available common resource (Sects. 7.10 and 7.13.3);

(h) the *Rings* acting in social systems as CAS (Sect. 7.11);

(i) the *Rings* that produce change in social systems; the PSC model of change (three-wheeled, three-speed) (Fig. 7.23) was presented (Sect. 7.12).

Chapter 8
The Magic Ring in Action: Organizations

None of our institutions exists by itself and as an end in itself. Every one is an organ of society and exists for the sake of society. Business is no exception. 'Free enterprise' cannot be justified as being good for business. It can only be justified as being good for society (Peter Drucker 1986, p. 33).

Although organizational learning occurs through individuals, it would be a mistake to conclude that organizational learning is nothing but the cumulative result of their members' learning. Organizations do not have brains, but they have cognitive systems and memories. . . Members come and go, and leadership changes, but organizations' memories preserve certain behaviors, mental maps, norms, and values over time (Bo Hedberg 1981, p. 6).

The repertoire of tools that top managers use. . . is determined by their experiences. A dominant logic can be seen as resulting from the reinforcement that results from doing the right things . . . they are positively reinforced by economic success (Prahalad and Bettis 1986, pp. 490–491).

The discipline of control entails observing reality by filtering it through the models produced by the simple theory of control discussed in Part 1 of this book. Chapter 6 took the point of view of the individual observing the environment he or she lives in; Chap. 7 examined the *Rings* that operate within populations, social systems, combinatory systems, and complex adaptive systems. Control systems also operate in production or consumption *organizations* which are unitary systems where the control is pervasive and essential. In particular, all production organizations are able to carry out their activities and survive over a long period of time precisely because they are preordained to undertake a constant control of the coordination of individuals and organs, the macro process they carry out as a single entity, and the network of micro processes of their organs. For this reason organizations must necessarily create control systems at different levels (which are interconnected) in order to regulate all aspects of their performance, the internal organizational micro processes, and the relations with the market and external environment. As unitary wholes, organizations must themselves be considered control systems, whose lengthy existence depends, however, also on lower level control systems pervading every part of their structure, their processes, and, in the final analysis, their existence. In particular, we shall examine organizations from

five points of view: (1) as autopoietic and teleonomic systems, (2) as vital systems, (3) as systems for efficient transformation, (4) as systems subject to management control, both entrepreneurial strategic and managerial operational, and, finally, (5) as cognitive and explorative systems. Once again we shall be able to see the action of the magic *Rings*, without which organizations could not exist but would instead become transformed into a group of agents operating in an individualistic and disorderly manner.

8.1   *Rings* That Allow Organizations to Exist as Autopoietic, Homeostatic, and Teleonomic Social Systems

It is now time to examine the action of the control systems we can observe, or design, in organizations, in particular *business* and *profit-oriented* production organizations.

Organizations are special *social systems* made up of individuals or *agents*, or groups of these, linked by a network of *stable organizational relations* or *structural constraints*—the "organization," in technical words—that specify, for each structural component, the following five elements: (1) a precise spatial and temporal *placement* in a hierarchy (topology), (2) a specialized *function* in relation to the macro function of the entire structure, (3) a specific *functionality* that delimits the admissible interactions of any agent with the others, (4) a set of standards of *functioning*, (5) a structure of control systems for the behavior and performance of agents and their processes.

The structural elements that conform to these five elements have no behavior or autonomous significance except in relation to the higher organizational level; they are *organs* of the system linked to the organization and to the functionality of all the organization's elements. Precisely by considering the organic structure of organizations as social systems, March and Simon proposed the analogy with biological systems, which, by definition, are organisms.

> Organizations are assemblages of interacting human beings and they are the largest assemblages in our society that have anything resembling a central coordinative system . . . the high specificity of structure and coordination within organizations—as contrasted with the diffuse and variable relations among organizations and among unorganized individuals—marks off the individual organization as a sociological unit comparable in significance to the individual organism in biology (March and Simon 1958, 4).

Through the network of stable relations among its organs, the organization is able to provide constraints to the behavior of its constituent agents in order to produce, at the various hierarchical levels, a network of coordinated and cooperative *processes* for the achievement of a common or an institutional goal; for this purpose the organization carries out the activities and operations needed to interact with its environment and achieve the objectives established by the governance in

order to satisfy the interests of parties outside the organization: the *stakeholders*. In fact, each organization, as a unitary entity, exists within a social system and is structurally coupled with a more or less vast environment, thereby creating a network with other organizations. Many groups, other than the internal agents, can benefit or be harmed by the organization; these groups taken together make up the *stakeholders* that define the limits to the freedom of action of the organization, carrying out an external control of the dynamic processes leading to the achievement of the institutional goal.

An organization is created from the desire of a "founder" to achieve some goal; the "agents" willingly participate in the organization based on their *specific individual motivations*—usually a monetary compensation and a career path—accepting to limit their potential in order to become (or be a part of) *organs*. These agents thus *formally recognize* the constraints imposed by the *organizational relationships* and by the structural *ties* that oblige them to become "structurally coupled" with other "agents," giving rise to *specialized, coordinated,* and *cooperative* behavior. Thus, these agents agree to become part of an organizational structure and accept the objectives, plans, rules, and responsibilities as constraints on their behavior *inside the organization*, on the condition that the latter allows them to satisfy their *specific individual motivations* (Cyert and March 1963). Organizations differ from social systems in many ways, the most evident of which is the fact that the agents participate voluntarily in the organization because of the advantages it brings them and can also voluntarily cease to collaborate with it. In social systems the agents in most cases are obliged to participate and cannot decide not to do so, except with limited exceptions.

The *agents* and *organs* they constitute can be considered *processors* that allow the structure to produce a *network of recursive micro processes*—financing, investment, production of value, disinvestment and reinvestment, reimbursements and refinancing, etc.—which tend to maintain and perpetuate themselves over time, continually adapting to the environment. Such *micro* processes form larger processes, managed by higher level organs which tend to recursively regenerate themselves over time, to the point of forming a *macro* process (attributable to the organization as a whole) which transcends the *micro* processes produced by the organs, in that it represents the *emerging* result of the *network* of micro processes. The macro process produces the organization's outputs, the qualitative-quantitative results perceived by the environment, which taken together form the *outcomes* directed at the stakeholders. As a result of this, each organization, considered as an *organizationally closed system*, can be observed at three different levels, as shown in Fig. 8.1.

Starting from the bottom, Fig. 8.1 represents:

1. The *level of the organizational structure* of the agents, with the organizational relations and the behavioral constraints (lower level in Fig. 8.1).
2. The *level of the network of processes* regarding the activities and operations the agents must carry out (middle level in Fig. 8.1); the vertical upward arrows from
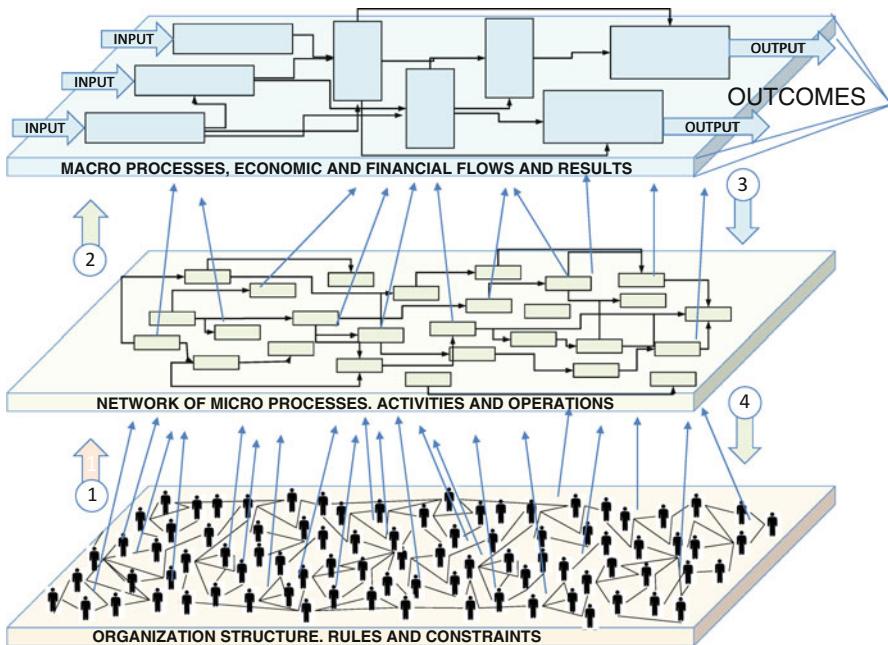
Fig. 8.1 The three interconnected levels of the organization

level 1 indicate that the network of processes is activated by the behavior of the agents and the organs in level 1.

3. The *level of the macro processes* that produce the financial and economic flows which allow the organization to achieve the results, outputs, and outcomes necessary to attain its institutional goals and to exist for a long time (upper level in Fig. 8.1); the vertical upward arrows from level 2 indicate that the macro processes that produce the organization's outcomes derive from the micro processes on level 2 and thus the behavior of the agents on level 1.

The three levels are thus interconnected by ascending and descending relations. The agents and organs in the organizational structure produce the activities and operations needed to carry out and maintain the *micro* processes (arrow①); the *micro* processes are necessary to carry out the *macro* processes which, given the inputs, produce the outputs, results, and outcomes (arrow②). The *macro* processes needed to achieve the results condition the composition of the network of micro processes (arrow③), which in turn conditions the organizational structure (arrow④).

Organizations survive as long as they can maintain their organizational constraints stable, replacing departing agents and inserting new ones that maintain the network of micro processes in order to activate the macro processes needed to achieve the goals. As long as they manage to repeat this cycle, organizations remain *vital*. In this sense organizations can, in all respects, be viewed as *autopoietic*
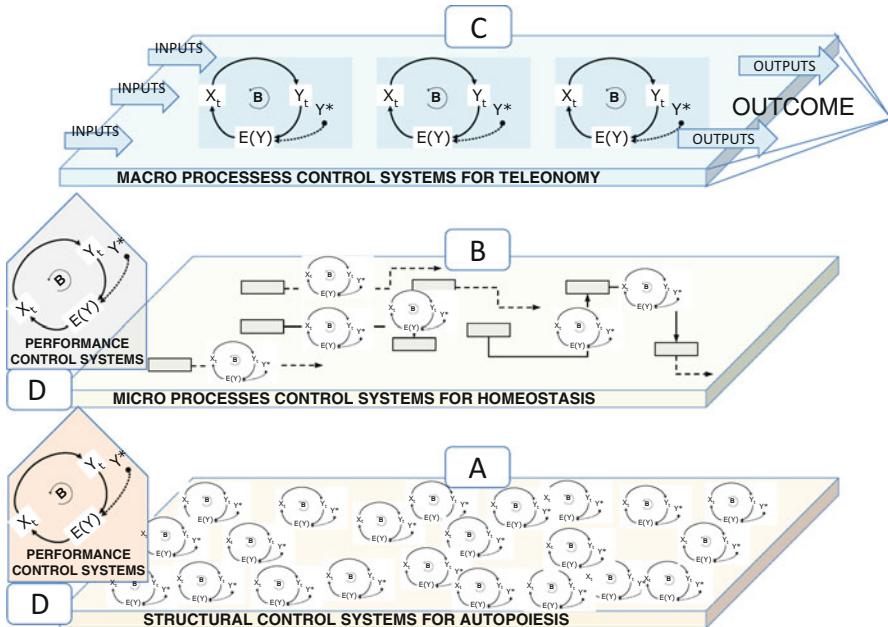
Fig. 8.2 The four levels of control systems for maintaining the organization as a unitary entity: [A] autopoiesis, [B] homeostasis, [C] teleonomy, [D] performance

machines, and thus *homeostatic* in line with Maturana and Varela's (Sect. 7.5) conception, in which organizations tend to remain in existence for a long time by continually regenerating the *processors* (organs) and the *network of processes* which form the "organizational fabric." And as we know already: "[. . .] *an autopoietic machine continuously generates and specifies its own organization through its operation as a system of production of its own components, and does this in the endless turnover of components under conditions of continuous pertur- bations and compensation of perturbations*" (Maturana and Varela 1980, p. 79).

If we accept this view, then it is clear that the existence of every organization, as a single entity, is conditioned by *four* levels of control systems which, as shown in the model in Fig. 8.2, act on the elements on the three structural levels in Fig. 8.1 in order to maintain the autopoietic networks:

(a) The control systems of the structural coupling of the individuals in the organi- zation based on the organizational constraints: I define these as the *Rings* needed for *autopoiesis*.
(b) The control systems for the efficiency, effectiveness, and equilibrium of the operations, activities, and micro processes that make up the autopoietic network the organization needs to produce the global macro process: I define these as the *Rings* necessary for *homeostasis*.
(c) The control systems to ensure the effectiveness of the macro processes that generate the inputs, outputs, results, and outcomes of the organization as a unit:

I define these as the *Rings* needed for the *teleonomy* of the organization, since they function to guarantee that the outputs, results, and outcomes respond to the demands of the stakeholders, which represent the maximum *vital constraints* for the achievement of the institutional objectives.

(d) The control systems for the effectiveness of the *performance* of the agents and organs, which form the structure (lower level in Fig. 8.1) needed to activate and maintain the network of micro and macro processes: I define these as the *Rings* needed for the performance of the agents, organs, and entire organization.

These four types of *Rings* are in line with the traditional four forms of organizational control: *personnel* control, *action* control, *results* control, and *performance* control (Merchant 1985). The *Rings* for *autopoiesis* (type A) form a network of *organizational* control systems aimed at producing an "alignment" among the agents in the organization and the various organs, so that their actions, functions, positions, tasks, and objectives will be *coordinated* and *cooperative over time*, which is a necessary condition for the long-lasting maintenance of the mutual *structural coupling* of agents and organs. I shall define this internal form of *control* of the structural coupling of agents and organs as a *structural self-control*, which is vital for any organization, and use the term *structural control systems* to refer to the *Rings* which achieve this; this form of control has also been defined as the control of *conformity* or bureaucratic control.

> Organizations require a certain amount of conformity as well as the integration of diverse activities. It is the function of control to bring about conformance to organizational requirements and achievement of the ultimate goals of the organization. The co-ordination and order created out of the diverse interests and potentially diffuse behaviors of members is largely a function of control [...] But regardless of how order is created, it requires the conformity of all or nearly all to organizational norms (Tannenbaum 1962, pp. 237–238).

The *structural self-control* also performs a *control* or a *surveillance of the position and behavior* of the various agents and organs, since each agent or organ in the structure has his or her/its own function to carry out within a specific field of action, according to shared standards, while respecting the external and internal norms based on a common ethical view.

> Cultural or normative controls which operate as alternatives to bureaucratic rules and direct supervision are new technologies of power developed within knowledge-intensive organizations. The mobility and extensive exposure of ... workers to a range of cross-cutting professional obligations coupled with the uncertainty and ambiguity of their tasks makes such control fragmented and contingent (McKinley and Starkey 1998, 10).

Structural self-control is made possible by the fact that the agents in the organization must accept the *responsibilities* that derive from their participation in the structure and respect the *authority* of the other agents and organs, which is established by the rules that discipline the structural links, as shown in Fig. 8.3.

The *structural coupling* control systems (lower level A in Fig. 8.2), as shown in Fig. 8.3, apply to every agent or organ (for example, A) in order to maintain or restore the coordination and cooperation between the position of A, the assigned tasks, and the
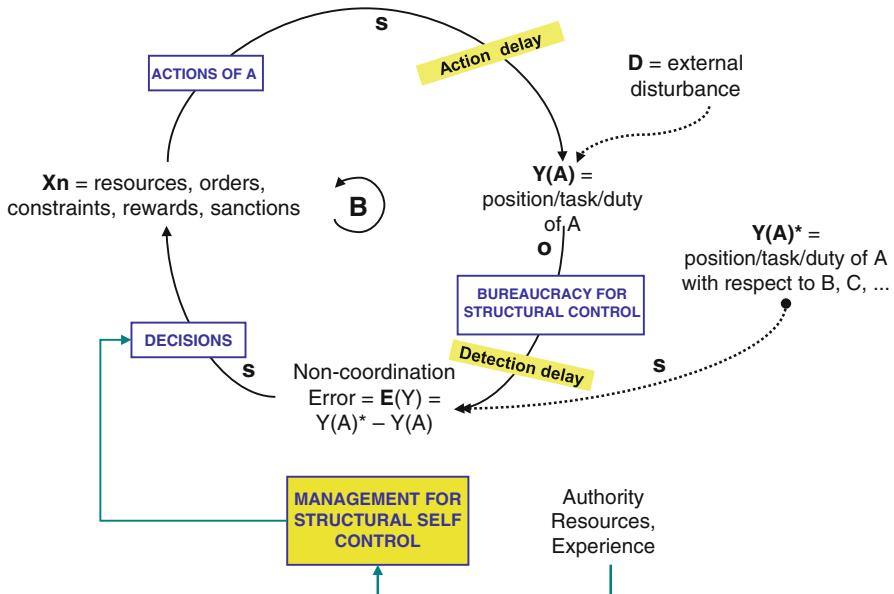
Fig. 8.3 Micro *structural control system* for the structural alignment of agents

procedure for the functions and processes with respect to those organizationally assigned. These systems use a variety of control levers, [Xn], such as resources, information, orders, constraints, incentives, persuasion, awards, punishments, and career possibilities, to eliminate the gaps, $E = Y^* - Y$, which are "symptoms of non-coordination" and the "absence of cooperation," between the position/task/objective of A, $Y(A)$, and the values required to maintain the *structural coupling* of A with B ($Y^* =$ position/task/objective of A to allow coordinated and cooperative behavior with respect to B), taking into account the external disturbances (D), which alter behavior.

It is immediately clear that a large number of *Rings* are required to guarantee the structural coupling of all the agents or organs in the organization (lower level A in Fig. 8.2). *Structural self-control* must be carried out by *steering systems*, generally multi-lever, that continually corrects errors in the structural coupling, taking into account the fact that each change in the behavior of an agent A determines a consequent change in the behavior of at least one agent B, thereby risking possible interference and oscillating behavior, according to the definition below of *structural coupling*:

Autopoietic systems may interact with each other under conditions that result in behavioral coupling. In this coupling, the autopoietic conduct of an organism A [individual or organ] becomes a source of deformation for an organism B [individual or organ], and the compensatory behavior of organism B acts, in turn, as a source of deformation of organism A, whose compensatory behavior acts again as a source of deformation of B, and so on recursively until the coupling is interrupted. In this manner, a chain of interlocked interactions develops such that, although in each interaction the conduct of each organism [individual or organ] is constitutively independent in its generation of the conduct of the
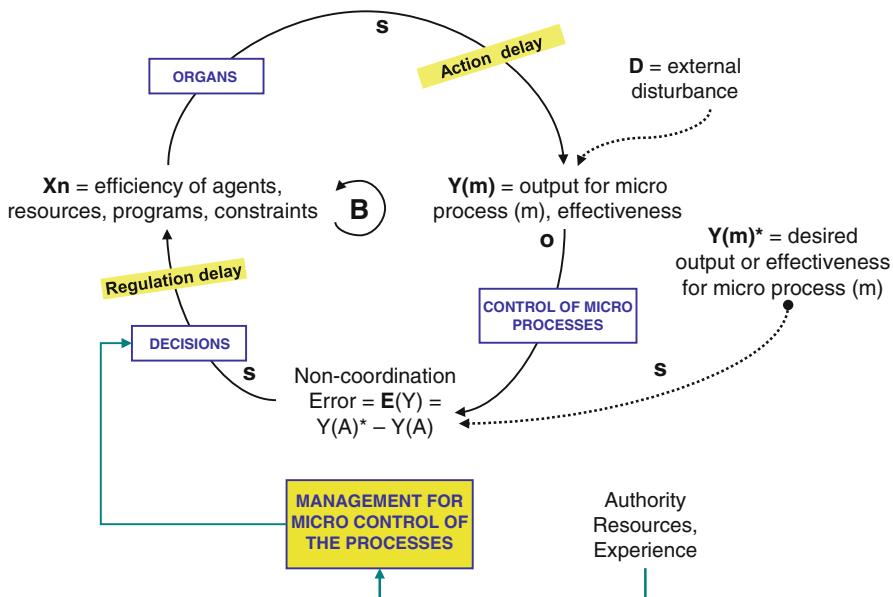
Fig. 8.4 Control systems for the control of *micro* processes

other, because it is internally determined by the structure of the behaving organism only,
it is for the other organism, while the chain lasts, a source of compensable deformations
which can be described as meaningful in the context of the coupled behavior (Maturana
and Varela 1980, 119–120).

Equally important is the function of the *Rings* that control the network of micro
processes necessary to ensure *homeostasis* (type B in Fig. 8.2). These processes
must guarantee that the *activities* and *operations* of the agents applied to the
network of micro processes conform to the established standards of behavior and
efficiency and produce the expected results in order for the higher level macro
processes to be produced, as shown in the model in Fig. 8.4.

Clearly this form of control is not aimed at guaranteeing the structural coupling
of agents (Fig. 8.3) but at ensuring that the behavior of the agents entails the
operationality and the efficiency required to produce the *outputs* of the processes
in which the agents are the processors. For this reason the control of the *network of
processes* is, in fact, the control of the operationality and behavioral efficiency of
the structure (agents and organs), and it seeks to align the *results* and *individual
performance* (agents and organs) to the standards of quality, efficiency, and pro-
ductivity set as constraints or objectives in order to ensure that the processes
themselves operate according to the functional standards for the higher level
network of macro processes.

A distinction is drawn between two modes of organizational control, one based on
personal surveillance, behavior control, and the other based on the measurement of outputs,
output control. A study of employees over five levels of hierarchy shows that the two
modes of control are not substitutes for each other, but are independent of each other.
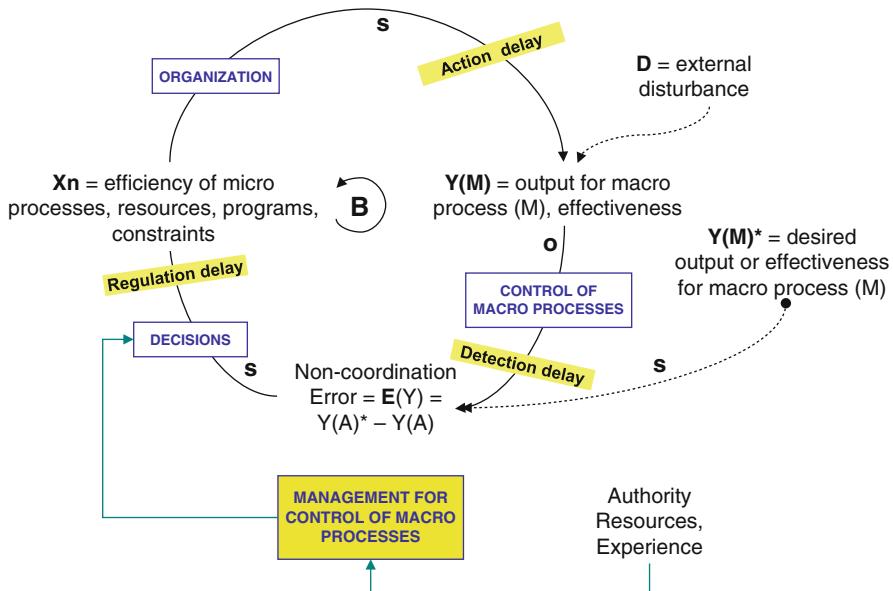
Fig. 8.5 Control systems for the control of *macro* processes

> The evidence suggests that output control occurs in response to a manager's need to provide legitimate evidence of performance, while behavior control is exerted when means-end relations are known and thus appropriate instruction possible (Ouchi and Maguire 1975, 559).

Along with the *Rings* (types A and B above) operating *within* the organization at the *micro* levels of the individuals and organs, control systems must also operate at the *macro* level (type C in Fig. 8.2), *on* the organization as a *unitary system*, to guide its complex activities toward the institutional objectives. In order to accomplish this, such control systems regulate the macro processes to produce the *flows, inputs, outputs, results,* and *outcomes* necessary to satisfy the outside stakeholders' expectations, as shown in Fig. 8.5.

This incessant activity regarding the *production of outcomes* and the *search for inputs* to maintain the organization in existence shows that organizations, in addition to being autopoietic and homeostatic systems, also produce *teleonomic* behavior to achieve the "existential project" for which they were created, according to Jacques Monod's definition in Sect. 7.7. I shall refer to these *control systems* as the *Rings for teleonomy*, since they drive the macro processes toward the attainment of the outputs necessary to achieve the institutional goals (Fig. 8.5).

We immediately perceive the importance of control systems of performance (type D in Fig. 8.2). The structure of the agents and organs is instrumental for the workings of the network of operational *micro* processes necessary for the functioning of the *teleonomic macro* processes. Thus the activity of the structure's components must conform to the *results* and *performance* standards required for the activities and processes to be carried out. The *Rings* for the *control of performance*
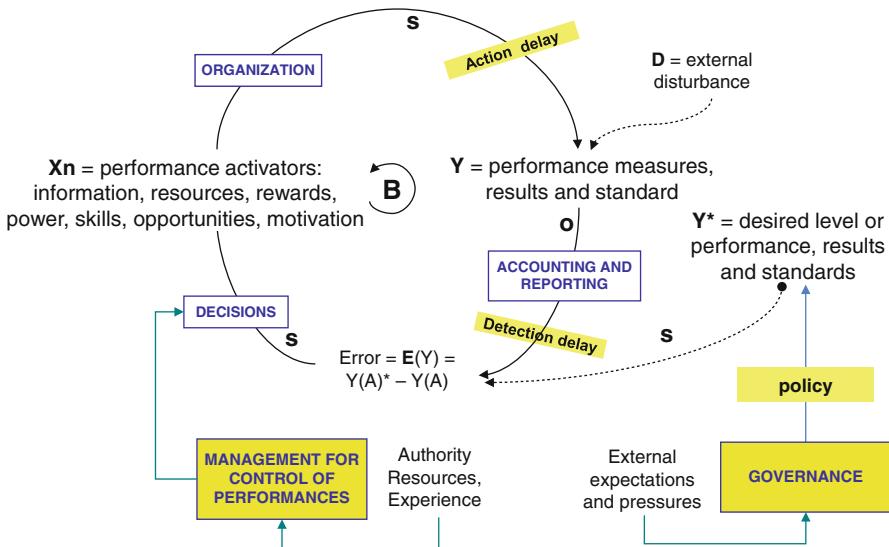
Fig. 8.6 Control systems for the control of *performance*

(shown in Fig. 8.6) should activate the levers needed to bring the individual performance to the levels required by the tasks and functions, since the organization, as a unitary entity, must satisfy the needs of the *stakeholders*, thereby producing outputs, results, and outcomes which meet expectations.

If it succeeds in efficiently structuring the four networks of *Rings*—for *autopoiesis, homeostasis, teleonomy,* and *performance* of the agents and organs—the organization, as a unit, will manage to maintain a stable structural coupling *with the environment* not only by reacting to *environmental* disturbances through its own system of *Rings* but also by searching the environment for disturbances (variables, stimuli, inputs) which are useful for autopoiesis (Zeleny and Hufford 1992; Mingers 1995, 2002) while fleeing from those which are harmful.

In this sense the long-lasting organization displays a cognitive behavior aimed at survival, and it can be viewed as a living system that reproduces itself over time, along the lines of Maturana and Varela's analysis, a part of whose quote from Sect. 6.7 states: "*If living systems are machines, that they are physical autopoietic machines is trivially obvious* [. . .] *However we deem the converse is also true: a physical system, if autopoietic, is living*" (Maturana and Varela 1980, 82).

The idea that the organization is a living system has been excellently described by Salvatore Vicari in a convincing book entitled *The Organization as a Living System* (Vicari 1991) and by Arie De Geus in his work *The Living Company: Habits for Survival in a Turbulent Business* (2002; see also 1997). De Geus clearly shows the importance of cognition and learning for an organization's teleonomy, especially large corporations, whose teleonomic activity can be interpreted only by assuming that the organization (company) is a living being and the decisions for organizational activities taken by this living being result from a learning process. It is not without significance that the *Forward* of this work was written by Peter Senge, who sums up

the reasons organizations must be viewed as living beings and not as simple machines. Among these reasons, I find the following quite convincing:

> Seeing a company as a machine implies that its actions are actually reactions to goals and decisions made by management. Seeing a company as a living being means that it has its own goals and its own capacity of autonomous action.
>
> Seeing a company as a machine implies that it will run down, unless it is rebuilt by management. Seeing a company as a living being means that it is capable of regenerating itself, of continuity as an identifiable entity beyond its present members.
>
> Seeing a company as a machine implies that its members are employees or, worse, "human resources", humans standing in reserve, waiting to be used. Seeing a company as a living being leads to seeing its members as human work communities.
>
> Finally, seeing a company as a machine implies that it learns only as the sum of the learning of its individual employees. Seeing a company as a living being means that it can learn as an entity, just as the theater troop, jazz ensemble, or championship sport team can actually learn as an entity. In this book Arie argues that *only* living beings can learn (Senge 1997, pp. IX–X).

Francisco Varela holds that non-biological entities should not be defined as living in the traditional sense:

> Frankly, I do not see how the definition of autopoiesis can be directly transposed to a variety of other situations, social systems for example. It seems to me that the kinds of relations that define units like a firm (Beer) [concern] units [which] are autonomous, but with an organizational closure that is characterizable in terms of relations such as instructions and linguistic agreement (Varela 1981, p. 38).

In this sense, *if observed from the outside* organizations, in their relations with their environment, are dynamic systems capable of self-regulation which themselves can be considered to be macro control systems, as shown in the model in Fig. 8.7. Each time there is a gap between the *outcomes* (Y*) the organization must provide its stakeholders to remain vital, and the *potential outcomes* (Y), *cognitive* and *operational levers*, are activated to search for the appropriate inputs and resources to produce the desired outputs and outcomes.

Whenever the organization is not able to produce a *macro* control to attain the desired *outcomes* for its stakeholders and to find the necessary *inputs*, it is incapable of reproducing its own organs (processors) and processes; the "organizational fabric" declines to a level such that the autopoietic network of vital processes is dismantled. The organization's *existence* comes to an end and the organization dies out, or changes in an irreversible manner, losing its identity. In this sense *cognition* and *learning* are necessary for the existential success of the organization; as Peter Senge has masterfully shown, the organization must become a *learning organization* as soon as possible (Senge 2006).

There is thus an evident strengthening loop between the *macro control*, which permits the long-lasting structural coupling between organization and environment, and the *micro controls*, which guarantee coordination, cooperation, and efficiency of the performances of the agents and organs in the structure.
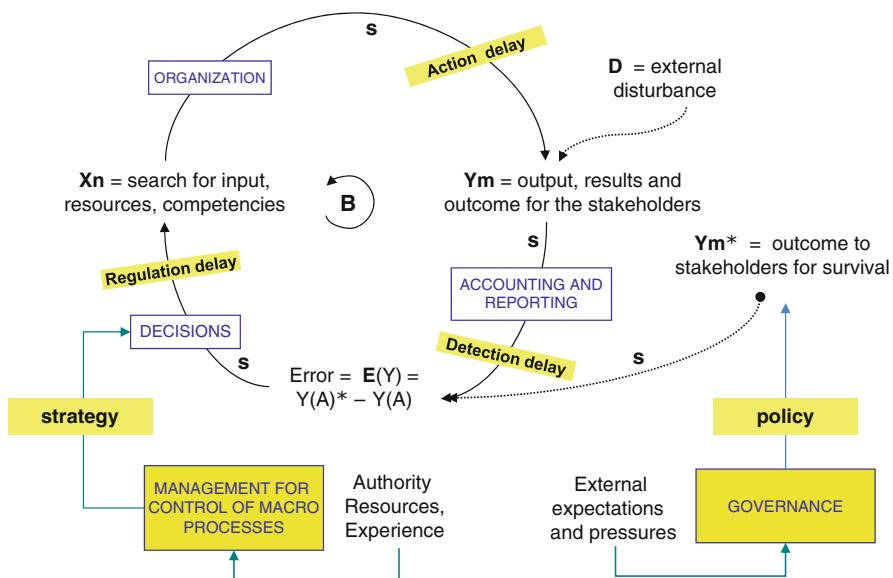
Fig. 8.7 The organization as a control system which regulates teleonomic behavior

8.2   Viable Systems View

The above considerations lead us to a general conclusion: organizations of any kind, in particular *production, business-oriented* ones (firms), can, as autopoietic systems coupled to the environment, have a long-lasting existence if (1) at the *macro level* they operate as unitary control systems, equipped with an internal *chain of control* produced by the organizational structure, capable of achieving the organizational outcomes, and (2) at the *structural level* they are composed of micro *Rings* that guarantee *autopoiesis, homeostasis*, and *teleonomy*. There are specific theories and models which allow us to represent organizations–enterprises in general, firms in particular, as unitary control systems. In this *section* I will start from Stafford Beer's model (Beer 1979, 1981), universally known as the Viable System Model, or VSM, which is particularly significant. The next *section* takes up a similar model, but one aimed at the objectives of control.

Directing the reader to Beer's book for a detailed description, here it is enough to mention that Beer directly interprets organizations and firms, when observed from the outside, as *cybernetic systems*, and thus as unitary control systems that include management directly in the chain of control. In order for organizations to have a long-lasting existence, management must continually modify their behavior based on changes in the objectives and constraints set by the stakeholders and by outside events. Thus, Beer believes that organizations must be viewed as *viable systems* which, through their structure, which is capable of learning and cognition, can

achieve an enduring structural coupling with the environment, continuing in this way to exist for a long time through continually adapting to the environment. To maintain the conditions for *viability*, organizations, as control systems, internally determine the policies and activate the levers and strategies needed to eliminate the negative effects from environmental disturbances during the course of their existence, disturbances which cannot be foreseen at the moment the system is designed and created. In his book *Brain of the Firm*, Beer provides a definition of viability:

> This book has been wholly about the viable system. There must be criteria of 'independent' viability, even though any system turns out to be embedded in a larger system and is never completely isolated, completely autonomous or completely free (Beer 1981, 226).

> The object is to construct a model of the organization of any viable system. The firm is something organic, which intends to survive—and that is why I call it a viable system (*ibidem*, 75).

In this sense, even the VSM, in line with the *autopoietic* viewpoint, represents organizations as cognitive, adaptive, long-lasting (though the exact length of time is not predefined) systems coupled to the environment through systems of communication and exchange. However, the VSM takes a *macro* and *external* perspective, according to which organizations are entities that survive thanks to their cognitive and control structure and are able to communicate with the environment and acquire information needed to define and achieve, through the coordination of the operational units, their institutional objectives. From this point of view, and adopting Varela's analysis, the VSM represents the typical model of a *system open* to the environment through its own operations and communication activities, though *closed* regarding the processes of adaptation to environmental dynamics, as synthetically shown in Fig. 8.8.

However, conceiving of organizations as autopoietic systems (Sect. 8.1) adopts a mainly *micro* point of view, considering mainly the logic of the self-production and homeostasis of the "organizational fabric" and of the autopoietic *network* of processes considered to be *organizationally closed*, in the sense they form an unvarying structure that the system maintains stable and in homeostatic equilibrium over time (Bednarz 1988). For Varela (1979, pp. 54–59), a system that must carry out internal processes {P} as a condition for its existence, as typically occurs in human organizations (Sect. 8.1), has *organizational closure* if (1) the relationships among the processes {P} comprise a network {T} and not a simple hierarchy; (2) the processes in {T} are mutually interdependent for their own generation and realization; and (3) the set of processes {P} constitute the system as a unity "recognizable in the space (domain) in which the processes exist" (Ibid., p. 55). Given these assumptions, the VSM appears to be a powerful instrument for understanding the *structural conditions* of vitality for the firm and at the same time can be interpreted using the same basic concepts adopted for autopoietic and cybernetic systems.

Beer's model is graphically very complex; however, in its essential schematic form it can be summarized as follows.

SS1: Operations: This represents the *operational processes* of the organization, which are carried out by distinct *operational units* that can be observed in both their
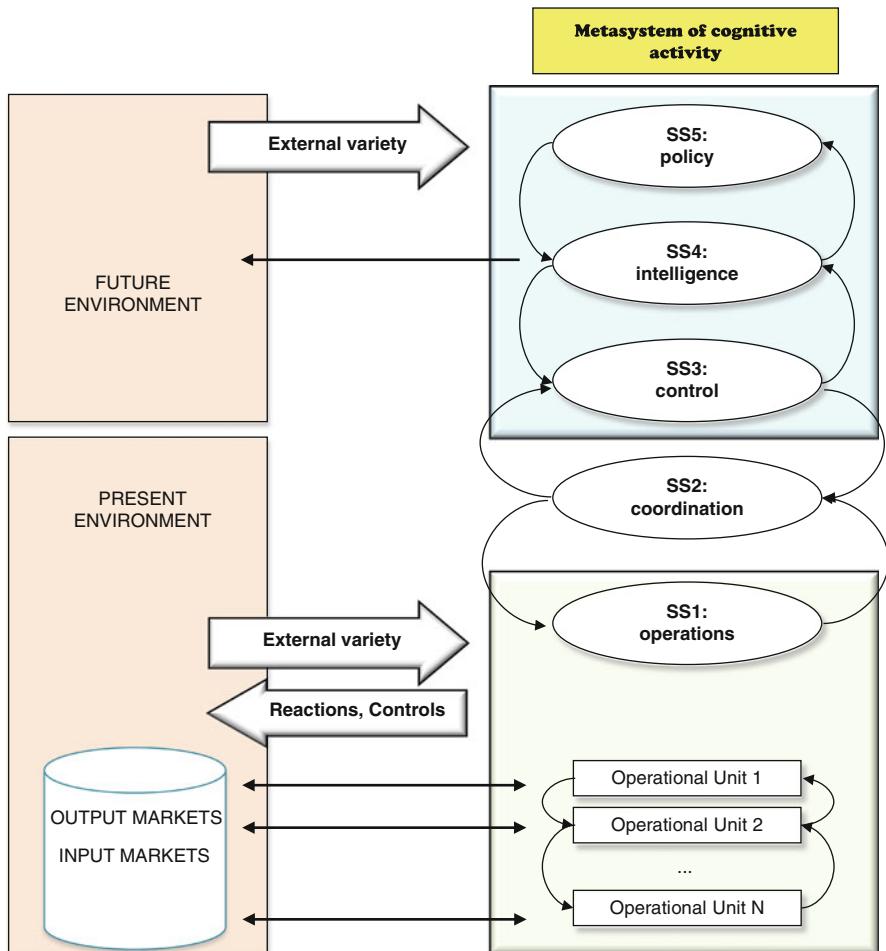
Fig. 8.8 A synthesis of the viable system model

vertical relations—aimed at the achievement of the objectives of vitality and survival of the entire organization—and their *horizontal relations* involving inter-exchanges with the environment, to which, based on the autopoietic view, they are *structurally coupled* in order to achieve their own objectives to ensure survival. Thus, the operational units which make up SS1 are control systems directed at specific and particular objectives and constraints, internal as well as external.

SS2: COORDINATION: Clearly, the operational units of SS1, since they employ common resources and can be in competition with the other units regarding objectives, are *interconnected* and normally *interfering*. As such they may be lacking in coordination and produce oscillatory dynamics in their local values which can cause inefficiency at the global system level. Precisely for this reason,

SS2 is assigned the task of coordinating the interconnected operational units by undertaking control actions to avoid the harmful effects from interference and the lack of coordination. For this role SS2 tries to optimize the *structural coupling* between the operational units of SS1, thus operating as a control system in a way similar to what is depicted in Fig. 8.3, coordinating the operational units in order that these can achieve their individual objectives while respecting the common constraints.

SS3: CONTROL: The operational units of SS1, insofar as they are coordinated by SS2, must be directed as a whole toward the achievement of the higher order objectives, which refer to the overall organization. SS3 is entrusted with this task. By assigning specific programs of action to the various units of SS1, as part of a general *programming* accompanied by a constant *monitoring*, SS3 directs its processes to achieving the final objectives of the system. The term "CONTROL" itself used by Beer to refer to this subsystem clearly indicates that SS3 is a typical control system based on a particular programming. As it is able to activate a *range* of control levers, even at various levels, SS3 is assigned the task of formulating *strategies* regarding the use of the levers for the various objectives. Nevertheless, SS3 cannot separate itself from subsystems 4 and 5, which together form a higher level subsystem that produces *cognitive activity* of the organization.

SS4: INTELLIGENCE OR RESEARCH REGARDING INFORMATION ON THE ENVIRONMENT: The survival capacity and the conditions of vitality do not depend solely on the present circumstances surrounding the environment, to which the organization is *structurally coupled*, but above all on future dynamics. It is thus vital for the organization to continually observe and forecast the "future environment" in order to enable SS3 to formulate programs of action and to adapt to these the units and activities of SS1. SS4 represents precisely the element of the *viable system* entrusted with proposing the vital objectives, based on predictable future scenarios, which are translated into programs of action, and with then identifying, gradually and in real time, the deviations from the programs, which are due to the actual environmental dynamics.

SS5: POLICY: To complete the VSM Beer has perceptively noted that the objectives set by SS4 could be mutually incompatible or in competition with one another and thus not always suitable to maintaining future conditions of vitality. Thus, he holds that every organization must provide for a higher level activity—a true governance—to identify the *hierarchy* of objectives set by SS4, whose pursuit through the activities of SS1, coordinated by SS2 and controlled by SS3, represents the condition for the survival over time of the organization as a vital system. The introduction of SS5 is provided with a *unitary management*, an *entrepreneurial* as well as *managerial capacity* to define the *policies* for the achievement of the vital objectives, explicitly recognizing the fact that organizations, as unitary entities, are *multi-objective* control systems and that, as such, *strategies* for the control levers employed by lower level subsystems do not suffice; instead, a careful policy evaluation and rational ordering of objectives are essential.
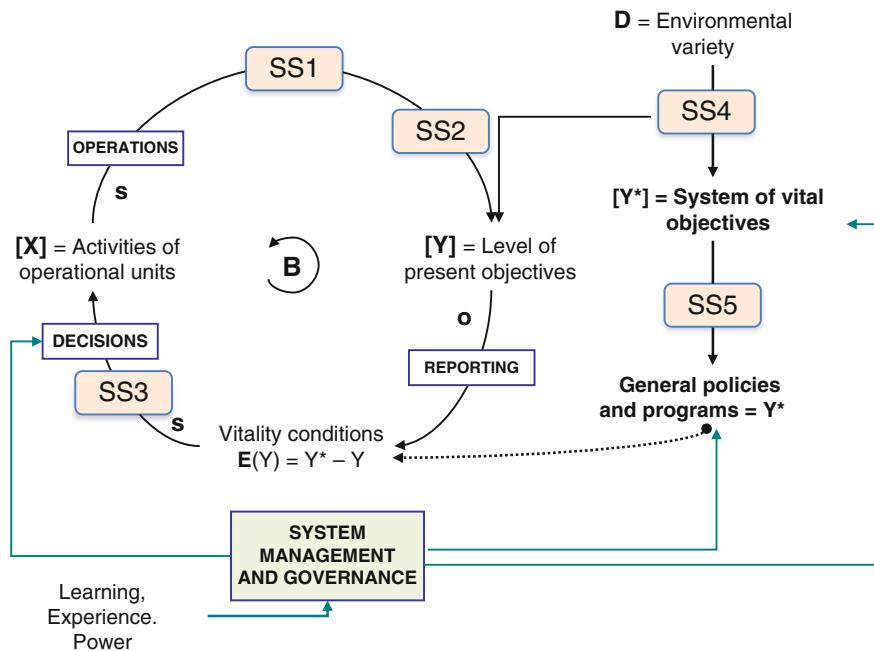
Fig. 8.9 The VSM as a multi-lever and multi-objective control system

In short, the VSM explicitly recognizes that every organization can be fully considered, in the joint action of its five subsystems, as a unitary, multi-lever, and multi-objective control system, as represented in Fig. 8.9. In this sense, the VSM (with the appropriate adjustments) can perfectly apply to organizations the heuristic schema used by the sociologist Talcott Parsons (1951, 1971) in order to interpret the general functioning of social systems, or society in general. This schema is called the "AGIL paradigm" ("scheme" or "pattern").

The AGIL paradigm is based, in fact, on the assumption that every social structure, as well as every organization considered as a social system with a stable closed structure (Sect. 7.5), seeks to remain vital and to guarantee its continuity with respect to its environment. To ensure its *vitality* every society must carry out four different functions or "functional imperatives" (the acronym AGIL derives from the initials of these functions).

Each of these imperatives corresponds to a particular subsystem of the social system.

A—*Adaptation* to the environment: in society the adaptation function involves the *economic subsystem*, that is, those structures that see to the production of the necessary resources (material goods, capital, knowledge and know-how, etc.) and the division of social labor, that is, the structuring and division of the tasks within the social structure to ensure adaptation.

G—*Goal attainment*, that is, the function needed for a society to attain its goals by seeking out the optimal use of resources to achieve society's main goals: In a

society the goal attainment function pertains to the *political subsystem*, which must define and pursue those goals held to have priority.

I—*Integration*, that is, the function that integrates and makes coherent the choices and actions of the structures making up the social system (subsystems, groups, roles, etc.), thereby guaranteeing the integration of mutual expectations and the rational use of resources: Integration thus concerns the normative structures, that is, those which formulate and impose respect of the laws and sanction any violations in order to maintain the social fabric.

L—*Latency pattern maintenance*, which is a sort of cultural template that guarantees the passing on of values, cultural models, and roles that structure a society, thereby ensuring that these are interiorized by all the members of the social system: This function falls to the educational structures (school, training programs, family, rites, systems of meaning, etc.) which are predetermined to allow each member of a society to acquire the behavioral and value models suited to his role and function. *One final note*. According to Beer, the viable system represents in all respects a holonic entity (Sect. 1.2), since every organization, while a complete unit, is in turn composed of smaller vital entities (operational units, departments, functions, divisions, etc.), with different levels of complexity, and at the same time part of a larger vital unit, as Beer clearly recognizes in the following theorem:

> *Recursive System Theorem*. In a recursive organizational structure, any viable system contains, and is contained in, a viable system. There is an alternative version of the Theorem as stated in *Brain of the Firm*, which expressed the same point from the opposite angle: 'if a viable system contains a viable system, then the organizational structure must be recursive' (Beer 1979, 118).

8.3   Organizations as Efficient Systems of Transformation

The preceding overview of Beer's viable system models refers to all organizations, independently of the nature of the processes they carry out. But what do production organizations and companies actually do to remain vital and effectively adapt to environmental changes?

To highlight the functions all organizations and firms must carry out to remain vital, I have proposed a specific model (in many respect parallel to the VSM) which considers organizations, and for-profit companies in particular, as *systems of transformation* which, in order to remain in existence over time, must carry out five interconnected vital *transformations*, each of which, operating with maximum efficiency, carries out a vital function similar to what is proposed in the VSM. Unlike the VSM, which represents organizations from the point of view of their *structural* synthesis, the Model of the Organization as an Efficient System of Transformation (MOEST) sees them from a *functional* viewpoint (Mella 2005b).

The MOEST can be applied both to *productive organizations* in general and to *enterprises* such as *business profit organizations* which, in *technically* producing goods and services, must also *economically* produce (and distribute) *value*. I shall

apply the model to *capitalistic enterprises* which, in order to produce, require capital inputs in the form of *equity* and *debt*, with the constraint (obligation) of maintaining or increasing the value; producing remuneration, dividends, and interest (output); and paying back the invested capital.

The MOEST—represented in Fig. 8.10—shows, above all, how each firm must necessarily carry out three efficient "technical" transformations, so defined because they concern the productive, economic, and financial functions instrumental in allowing the organization to maintain its vitality in order to satisfy, in particular, the needs of its stakeholders and shareholders.

The complete model is very detailed; however, a synthesis would be useful to better understand the logic behind it. Hereafter referring to the MOEST in Fig. 8.10, I will begin with the last three transformations, which I have defined as "technical" to distinguish them from the first two, which I define as "cognitive."

TR1-P: PHYSICAL PRODUCTIVE TRANSFORMATION, through which physical volumes of productive factors (F) are transformed into volumes of finished products (P): This is typically a transformation of utility: through the *operational units* that make up the *productive system*, the factors of production, having a given utility, are acquired (inputs) and transformed into products with a given level of quality capable of producing greater utility, which is sold to clients (outputs). The efficiency of the productive transformation is measured by two indicators: (1) *productivity* (π), understood as the capacity of the transformation to generate the maximum production output with the minimum inputs (consumption) of factors ($\pi = P/F$), and (2) *quality*, understood as the maximization of the *use function* (qlP) of the products.

TR2-E: ECONOMIC OR MARKET TRANSFORMATION, through which the organization, by activating a network of *economic* units and processes, businesses, and investment, tries to increase the value of the productive factors F (economic output), purchased at the input prices (pF), by using them to obtain production P that can be sold at remunerative prices (pP), thereby obtaining revenue (economic input): From the difference between revenues ($R = P \times pP$) and costs ($C = F \times pF$) the firm produces the *operating income* (operating profit) or EBIT: (economic inputs – economic outputs) = ($R - C$) = EBIT, which is the synthetic value of the TR2-E. The function of this transformation is to negotiate as "advantageously" as possible the prices of the production and of the factors in order to produce the *maximum* revenues (input) with the minimum costs (output). The *economic efficiency* of this transformation, understood as the capacity to cover cost flows with revenue flows over a sufficiently long period of time, can be quantified by three main indicators: (1) the EBIT = $R - C$; (2) the ratio between EBIT and cost of production, that is, by the *return on cost* (roc); or (3) the ratio between revenues and costs (e = economic efficiency index) or its inverse ($1/e$ = cost/revenue ratio). We can write the economic efficiency index in the following form:

$$e = \frac{P \times pP}{F \times pF} = \pi \frac{pP}{pF} \tag{8.1}$$
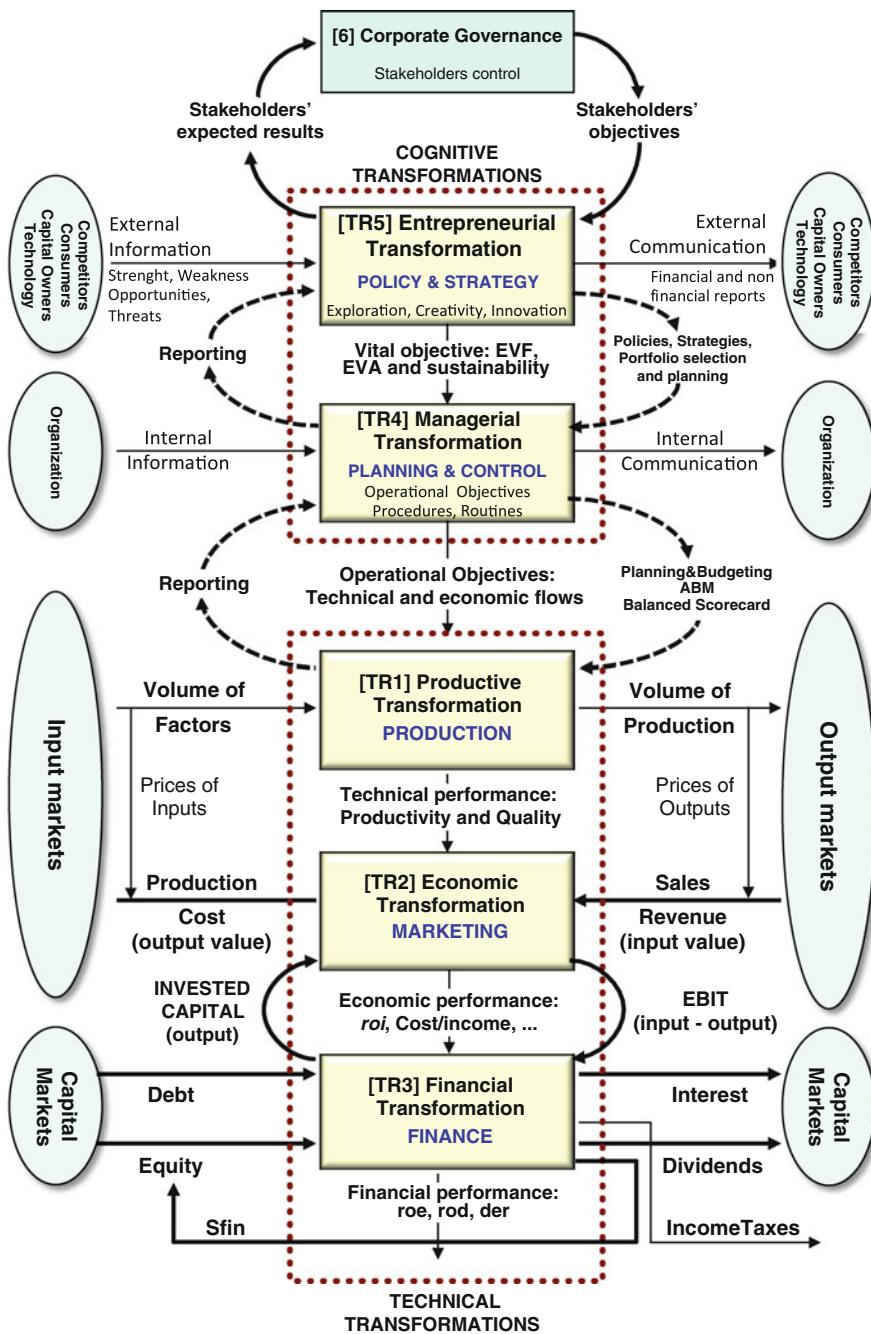
Fig. 8.10 The MOEST in brief (*source*: Mella 2005b)

We immediately observe that economic efficiency depends on productivity, which reflects productive efficiency, and on the ratio between the *average selling price* and the *average unit cost of production*, which represents *market efficiency*.

TR3-F: FINANCIAL TRANSFORMATION, through which the firm, in setting up its businesses, transforms the *acquired* capital (input)—in the form of equity (E) (underwritings, equity, stocks) or debt (D) (bonds, loans, and various forms of debt financing)—into *invested* capital ($I = E + D$), which is *used* (output) to finance TR2-E: By employing the invested capital (I) in TR2-E, the economic transformation is able to produce EBIT = R – C, which is then transformed into *flows* of interest payments (Int) for D and in dividends (Div) for E. Not all the EBT = (EBIT – I) becomes dividends, since a portion is assigned to taxes (Tax) and a portion is kept as self-financing (Sfin), which goes to increase equity for the growth of the firm. Only the residual amount is distributed as dividends. Thus, the financial transformation is typically a transformation of the risk of the capital investment in the hope of obtaining an adequate return. The efficiency of TR3-F is measured by the *yield on the capital*, which, as a first approximation, is quantified as the ratio between the *average return* on capital and the average capital amount, with reference to a given period. In detail, the main indicators of *financial performance* are (1) the return on equity (roe = EBT/E); (2) the return on debt (rod = I/D), taking into account the debt/equity ratio (der = D/E); and (3) the return on invested capital (roi = EBIT/I). Note that roi, though it is an indicator of the overall financial efficiency of TR3-F, also expresses the economic efficiency of TR2-E. Thus, roi summarizes the efficiency of all three transformations.

A necessary condition for the firm to activate the first three "technical" transformations is that two "cognitive" transformations also be carried out: the *entrepreneurial* (TR5-I) and *managerial* (TR4-M) transformations, whose function is to "control" the "technical" transformations. Though TR5-I is at the top of the model in Fig. 8.10, it is a preliminary requirement for TR4-M and for coherence sake will be examined first.

TR5-I: ENTREPRENEURIAL TRANSFORMATION, whose function is to monitor the *present* and *future* environments in order to (a) identify the survival conditions and define the *maximum objectives* that will guarantee an enduring vitality; (b) decide which *entrepreneurial policies* have priority in terms of the vital objectives; and (c) for each objective establish the *entrepreneurial strategies* to order the most effective control levers. To carry out this function, TR5-I produces a continual transformation of *information* and *forecasts* into *strategic decisions*, preparing the long-term plans and programs and designing the *management control systems* that give rise to and regulate the three other transformations for the achievement of the quality, productivity, economic efficiency, and profitability objectives. I have referred to the transformation as "entrepreneurial," as it produces to the maximum extent possible the conceptual, creative, and innovative activities that characterize the entrepreneurial function. Through TR5-I the firm, as a whole, can be viewed as a system in which present and future information can be transformed, by means of the

economic calculation, into *political* and *strategic decisions* as well as into *regulations* for the purpose of achieving the vital objectives; the firm adapts itself for this purpose to the environment in order to carry out the three efficient "technical" transformations. The maximum vital objective in profit organizations is the production of value for shareholders, who normally represent the most relevant component of the stakeholders. The production of adequate levels of *economic value of the firm* (EVF), from which shareholder value derives, and the maintenance of the conditions of *sustainability* represent the maximum objectives imposed by TR5-I on TR4-M. However, the entrepreneurial transformation is, in turn, controlled by the stakeholders, who activate the corporate governance and set the maximum institutional objectives and the environmental restrictions for the survival of the organization as a vital system.

TR4-M: MANAGERIAL TRANSFORMATION, which has a *control* function in order to allow the "technical" transformations to achieve, with the maximum efficiency, the *policies* and *strategies* decided by the entrepreneurial transformation: TR4-M undertakes five sub-functions: (1) it divides the vital objectives determined by TR5-I into operational objectives to be assigned to the organs (functions) and operational units; (2) it divides the *overall* entrepreneurial strategies drawn up by TR5-I into *functional* and *operational* strategies, which it assigns to the organs and operational units that carry out the "technical" transformations, and as a consequence (3) it draws up the operational programs and budgets that serve as the *operational objectives* for the control systems, which are required to achieve the maximum level of productive, economic, and financial efficiency; (4) it carries out the *managerial coordination* of the organs, operational units, and members of the organization that together represent the engines of the "technical" transformations; (5) it decides on the *operational regulations* which oblige the controlled units (organs, units, and individuals) to undertake the necessary actions to achieve the objectives (model in Fig. 8.2).

For convenience sake, I shall present in more concise form the system of the (macro) variables and of the basic values that regulate the MOEST to allow the reader to better understand the model in Fig. 8.10. The system of economic and financial values that represent the dynamics of the organization for any particular period can be summed up by the following balance sheet relations:

$$
\begin{cases}
L + I = D + E & \text{[financial position]} \\
C + \text{EBIT} = R & \text{[economic position]} \\
R = P \times pP & \text{[input of values]} \\
C = F \times pFP & \text{[output of value]} \\
\text{EBIT} = R - C & \text{[revenue production]} \\
\text{EBIT} = \text{EBIT} - \text{Tax} & \text{[earnings before taxes]} \\
\text{EBIT} - \text{Int} - \text{Tax} = \text{Div} + \text{Sfin} & \text{[revenue distribution]}
\end{cases}
\tag{8.2}
$$

In the above relations, L indicates liquidity (cash + receivables − trade payables), I is the invested capital, D and E represent debts and equity, Int = $(D\ i\%)$ represents the interest paid on D, Tax = [(EBIT − Int) tax%] denotes taxes, and Div and Sfin indicate the dividends and the self-financing provisions.

The system of performance indicators can be summarized as follows:

$$
\begin{cases}
e = \dfrac{R}{C} = \pi\dfrac{pP}{pF} & [\text{economic efficiency}] \\[2ex]
\pi = \dfrac{p}{F} & [\text{productivity} = \text{production efficiency}] \\[2ex]
f = \dfrac{F}{P} = \dfrac{1}{\pi} & [\text{unit requirement of factors}] \\[2ex]
cP = \dfrac{C}{F} & [\text{unit cost of production}] \\[2ex]
roi = \dfrac{\text{EBIT}}{I = E + D} & [\text{economic and financial efficienty}] \\[2ex]
roe = \dfrac{\text{EBT}}{E} & [\text{return of equity}] \\[2ex]
rod = \dfrac{\text{Int}}{D} & [\text{return on debt}]
\end{cases}
\tag{8.3}
$$

Figure 8.10, even if in extremely concise form, clearly shows that, even *functionally speaking*, every production organization must be viewed as a system of control systems that aims to achieve the objectives necessary for its survival. In fact, the MOEST includes three main forms of control: (1) stakeholder control (corporate governance in a broad sense) over the entrepreneurial transformation; (2) the strategic control of the entrepreneurial transformation, indicated in Fig. 8.11; and (3) the operational control of the managerial transformation, represented in the model in Fig. 8.12.

Figure 8.11 shows that the *entrepreneurial* transformation, TR5-I, identifies or receives from the governance the *vital objectives* for survival and determines the *policies* and *general programs* that become the *strategic objectives* the *managerial* transformation, TR4-M, must achieve through the control systems (normally defined as *strategic*), which act at the business and general function levels. Figure 8.12 indicates that the *managerial* transformation, TR4-M, translates the strategic objectives into *operational objectives*, to be achieved by means of a planning and budgeting program which is necessary for the *operational control* system to produce the necessary strategy to activate the available levers.

The *entrepreneurial* transformation, TR5-I, is, in turn, subject to an *institutional control* at an even higher level, carried out by the stakeholders, who represent the CORPORATE GOVERNANCE; the amount of control TR5-I has in the organization depends, in fact, on the limits set by the governance. The models in Fig. 8.11 and 8.12 illustrate
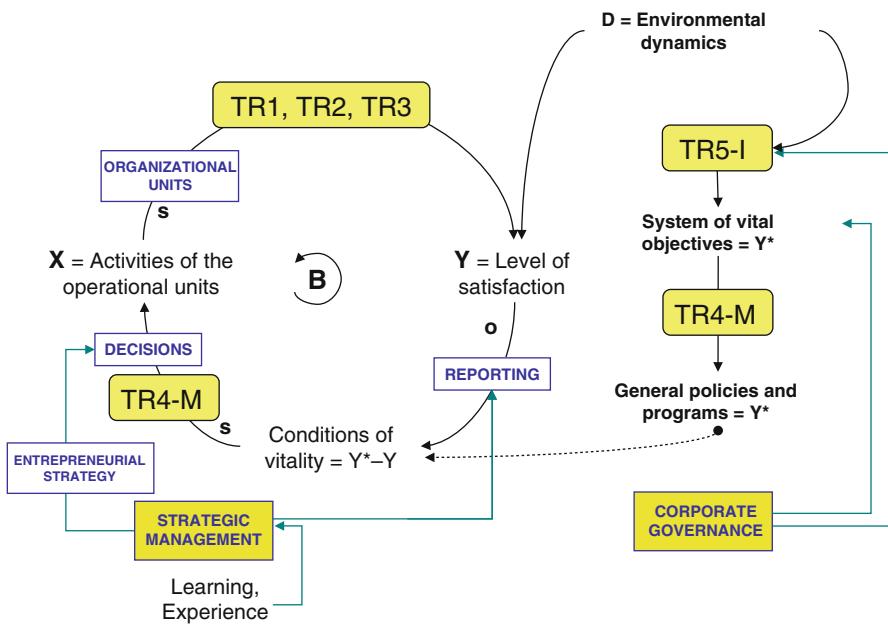
Fig. 8.11 The MOEST as a control system based on policies and strategies
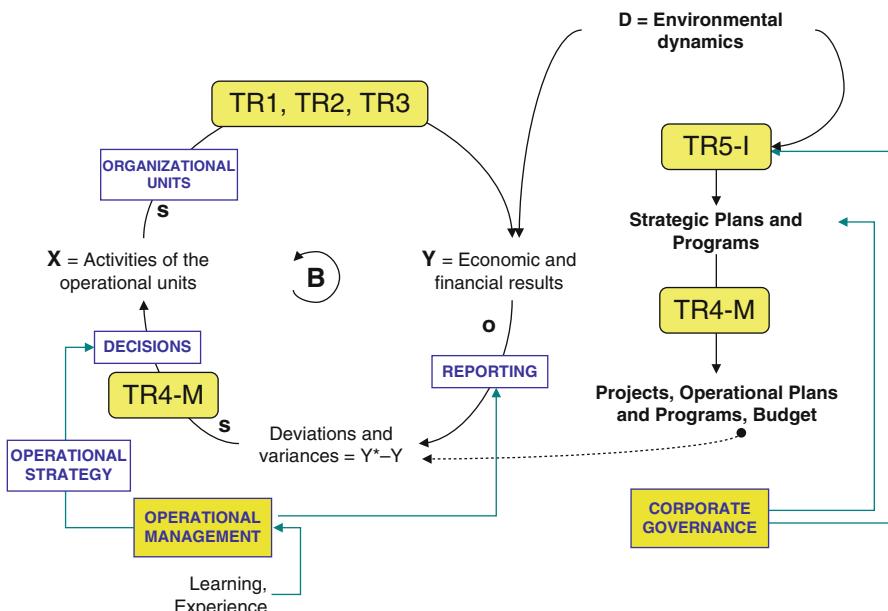
Fig. 8.12 The MOEST as a control system based on operational plans and budgets

the role of the three technical transformations in implementing the control in order to achieve the vital objectives.

The operational units of SS1 as described by Beer (Fig. 8.8) correspond to the units that carry out the "technical" transformations of the MOEST. The "cognitive" transformations of the MOEST, both entrepreneurial and managerial, perfectly correspond to the activities assigned to the other four subsystems of the VSM.

In this sense the MOEST is in line with Talcott Parsons' view of the organization. Parsons distinguished between three organizational levels, each characterized by its own level of organizational control: the *technical*, the *managerial*, and the *institutional* levels. The *technical* level includes those activities necessary to achieve the organizational goals. At this level, organizational control involves the operational activities necessary to achieve the organizational objectives. The *managerial* level is entrusted with the administration of the organization; at this level the organizational control must see to the obtainment of the necessary resources for the technical level and for the satisfaction of the customers and users. The institutional level must align all the organizational goals with the aims of the stakeholders; at this level the organizational control must maintain the identification and the uniformity between the organizational and social objectives.

There are two differences between Beer's model and the MOEST. Above all, the latter explicitly presents the three different levels of control: *institutional, strategic*, and *operational*, each of which pursues objectives different in scope. Secondly, it emphasizes the possibility of creating a precise system of *performance indicators* represented by the efficiency indicators: productivity, economic efficiency, and profitability, indicated in (8.3) (in Fig. 8.10 these are only mentioned), all of which can be linked from the production of value point of view.

The MOEST shows us that the function of control systems, at the institutional, strategic, and operational levels, is to guarantee that the typical transformations achieve the performance objectives regarding value, economic results, and profitability throughout the organization.

Observing the MOEST in Fig. 8.10, we can deduce the fact that the *profit organization's autopoiesis* depends on its ability to:

1. Create a dynamic portfolio of businesses along a virtuous path from "question mark" to "cash cow" businesses (following the well-known Boston Consulting Group matrix; 1970) by means of an effective *entrepreneurial function*.
2. Achieve the maximum exploitation of the present market and expand toward new markets in order to increase its production volume, P, and increase as much as possible the selling price, pP, through an efficient *marketing function*.
3. Contract the unit factor requirements while expanding the *quality* of products by means of an efficient production function, thereby increasing *productivity*.
4. Reduce the average factor prices through an efficient supply function which searches for supply markets where the factors have a higher quality and lower purchase prices.

The *autopoiesis of the capitalistic enterprise* is based on its capacity to regenerate its *financial* and *economic* circuits, which are made up of the value flows

represented by the bold arrows in TR2-E and TR3-F in Fig. 8.10. The financial circuit is renewed if the capitalist firm succeeds in acquiring and preserving its invested capital (I)—which is necessary for structural investments—by means of an adequate *financial leverage* (debt/equity ratio, or *der*); but this requires that the suppliers of both debt and equity financial capital receive a fair remuneration, defined as one at least equal to their opportunity cost.

8.4   From MOEST to Management Control and Performance Management

In the preceding *sections* I have observed that, no matter how we choose to view them, production organizations, firms in particular, are control systems that, through their particular structure, succeed in remaining *living* and *vital* by adapting themselves to a changing environment, recreating the "organizational fabric" (*micro* level), identifying their vital objectives (*macro* level), and trying to achieve these through control systems of varying size operating at all organizational levels (Fig. 8.2) and regulated by the respective managers. For convenience sake, I shall define *organizational control* as the network of all control systems which, operating at different levels (individuals, organs, functions, departments, businesses, variously defined operational units, etc.), *guide* the organization to produce the maximum *efficiency* in its operations and processes and to achieve the maximum *performance* in terms of productivity, quality, economic efficiency, and profitability, in an ever-changing environment and respecting the limits set by the market and the stakeholders. The network of these *Rings*, which ensure efficiency and allow the operational objectives to be achieved, is normally defined as a *management control system*; this network, as a whole, results in the *management control* of the organizational performance of the firm as a unit.

This conclusion is, in fact, not new to Management Science, Organization Science, and Management Control Theory; in point of fact, these fields have a strong tradition of research in the disciplines and techniques of management control, which Robert Mocker defines as follows:

> Management control can be defined as a systematic effort by business management to compare performance to predetermined standards, plans, or objectives in order to determine whether performance is in line with these standards and presumably in order to take any remedial action required to see that human and other corporate resources are being used in the most effective and efficient way possible in achieving corporate objectives (Mockler 1970, p. 14).

Very concisely, keeping in mind the general model of a control system (Fig. 2.10), we can conceive of management control as the vast process that specifies and operationalizes the efficiency and the performance objectives of the organization and *activates* all the "apparatuses" of the control systems needed to achieve the objectives: the *effectors*, which correspond to the operational units; the *detectors*,

which correspond to the monitoring, reporting, and performance-measuring centers; and the *regulators*, which correspond to the decision-making and responsibility units. Stafford Beer clearly illustrates this control system management role in organizations when he states that "*If cybernetics is the science of control, management is the profession of control*" (Beer 1966, p. 54). Even more clearly, in outlining the contribution of his book *Brain of the Firm* he states: "*This book is entirely concerned with the contribution which cybernetics, the science of control, can make to management, the profession of control*" (Beer 1979, p. 17).

It is not possible in this context to mention the thoughts of the many important management control scholars, not only due to the vastness of the bibliography but also because each author obviously follows his or her own conception of control and control system without, however, expressing this with an explicit definition. Nevertheless, it is useful to point out that, already in the 1960s, Robert Anthony presented a viewpoint of management control that reflected the different transformations indicated by the MOEST; in fact, he distinguished three types, or areas, of control systems: strategic planning, management control, and operational or task control.

> [Strategic planning is] the process of deciding upon objectives, on changes in these objectives, on the resources used to attain these objectives, and on the policies that are to govern the acquisition, use and disposition of these resources.
>
> [Management control is] the process by which managers assure that resources are obtained and used effectively and efficiently in the accomplishment of the organization's objectives.
>
> [Operational or task control is] the process of assuring that specific tasks are carried out effectively and efficiently (Anthony 1965, p. 16).

Referring to Fig. 8.10, it is clear that *strategic planning* is carried out by the *entrepreneurial* transformation in defining the time frame for the vital objectives indicated by the stakeholders and the overall strategy that indicates the strategic orientation for achieving these; that *management control* is carried out by the managerial transformation, which translates the strategic plans into specific control instruments for the operations of the "technical" transformations (Flamholtz 1996); and, finally, that operational control concerns the controls at the organizational structure level, which corresponds to the lower level in Fig. 8.2.

It is also useful to mention that the "management" concept itself has evolved to also include functions that differ from those related to the regulation of control systems. Already in 1916, Henri Fayol, in his book *Administration Industrielle et Generale* (translated in 1949 with the title *General and Industrial Management*), extended the concept of management to all activities of a firm, that is, to all the transformations indicated in the MOEST: "*technical, commercial, financial, security, accounting, managerial activities.*" However, he also emphasized the control activities that characterize the organizational structure of the agents and organs; in fact, he stated that "*To manage is to forecast and plan, to organize, to command, to coordinate and to control,*" indicating the well-known 14 *principles of* management (Kramer 2010). In 1975, William Newman, in spreading the concept of control systems, though with his own system of symbols, clearly demonstrated the essential and pervasive nature of management control of processes, directly underscoring the

relation between actions and objectives while stating that "*control is one of the basic phases of managing along with planning, organizing and leading.*" In his impressive book entitled *Management*, Peter Drucker emphasized above all the difficulties in appropriately using the term management while pointing out the empirical nature of management:

> The word "management" is a singularly difficult one. It is, in the first place, specifically American and can hardly be translated into any other language . . . It denotes a function but also the people who discharge it. It denotes a social position and rank but also a discipline and field of study (Drucker 1986, p. 5).

> Management is a practice rather than a science. In this, it is comparable to medicine, law, and engineering. It is not knowledge but performance. Furthermore, it is not the application of common sense, or leadership, let alone financial manipulation. Its practice is based on knowledge and on responsibility (*ibidem*, p. 17).

He also recognizes that management operates at all levels, even that regarding entrepreneurial transformation:

> For three-quarters of a century management has meant primarily managing the established, going business. Entrepreneurship and innovation, while mentioned in many management books, were not seen as central from 1900 till today. From now on, management will have to concern itself more and more with creating the new in addition to optimizing the already existing. Managers will have to become entrepreneurs, will have to learn to build and manage innovative organizations (Drucker 1973, p. 31).

It should be clear that management control refers above all to the typical actions of the managers—the *performance makers* in the firm, according to the MOEST model—which involve setting objectives for the control of processes, defining the organizational actions, identifying gaps regarding their achievement, modifying the organizational activities themselves, implicitly assuming that the control is supported by the actions of the agents and organs in the organization (the effectors in the control systems); and actions, agents, and organs which themselves must be controlled.

The network of control systems activated by management control specifically to carry out the technical, economic, and financial transformations constitutes, as a whole, the *management control system*. This includes a system for the production of information and organizational rules and procedures needed to activate, maintain, and improve organizational behavior, which is achieved by acting on individual behavior, as set out in the following definition by Robert Simons and David Otley.

> Control in organizations is achieved in many ways, ranging from direct surveillance to feedback systems to social and cultural controls [. . .] Management control systems [are] the levers managers use to transmit and process information within organizations. For the discussion to follow, I adopt the following definition of management control systems: *management control systems are the formal, information-based routines and procedures managers use to maintain or alter patterns in organizational activities* [. . .] These information-based systems become control systems when they are used to maintain or alter patterns in organizational activities (Simons 1995, p. 5).

> Management control systems provide information that is intended to be used by managers in performing their jobs to assist organizations in developing and maintaining viable patterns of behavior (Otley 1999, p. 364).

This definition clearly emphasizes the role of the manager as described by Beer and as it appears in the control system model (Fig. 3.1): determine the performance objectives; obtain information from the performance measurement systems, which detect deviations from the established performance objectives; make decisions regarding the control in order to decide on a strategy; and send information on the levers activated by the effectors (agents or organs), which are the true passive subjects of the control. Management control systems are thus vital for maintaining the organizations as viable systems, since they continually adapt the organization to the environment to continually satisfy the objectives set by the stakeholders.

> In broad terms, a Management Control System is designed to help an organization adapt to the environment in which it is set and to deliver the key results desired by the stakeholder groups, most frequently concentrating upon shareholders in commercial enterprises. Managers implement controls, or sets of controls, to help attain these results and to protect against the threats to the achievement of good performance (Merchant & Otley 2007, p. 785).

Since for every activity the performance is related to the results, but the latter depend on behavior, it is necessary to undertake not only a control of the operations but also a control on the *motivations* behind the behavior which impacts the *performance* of the operational units as well as on the *emotions* of the personnel which influence decisions and the standards of behavior. This form of control represents the typical focus of *performance management* techniques, a term introduced by Aubrey Daniels in 1970 with this definition: "*Performance is the sum of behavior and results, and cannot be viewed as independent of either component. It is an outcome of effective management*" (Daniels and Daniels 2004). To measure and improve the performance of individual and group behavior it is necessary to link them to systems of rewards or incentives (control levers) (Fig. 8.6). The more instrumental operational concept of *performance management system* derives from that of *performance management*. The former represents the set of all control systems and their networks which allow the organization, its organs, and agents to achieve and improve their performance (for an ample bibliographical review see Chiara Demartini 2013).

> We view PMSs as the evolving formal and informal mechanisms, processes, systems, and networks used by organizations for conveying the key objectives and goals elicited by management, for assisting the strategic process and ongoing management through analysis, planning, measurement, control, rewarding, and broadly managing performance, and for supporting and facilitating organizational learning and change (Ferreira and Otley 2009, p. 264).

This concept is by now widespread in all types of organization. For example, from a link ("Performance Management") on its website, Berkeley University in California clearly indicates this concept with the following definition:

> The University of California, Berkeley is committed to a performance management system that fosters and rewards excellent performance.
>
>   Effective performance management aligns the efforts of supervisors and employees with departmental and campus goals, promotes consistency in performance assessment, motivates all employees to perform at their best, and is conducted with fairness and transparency.

> The employee, the supervisor, and the University are critical members of a partnership that ensures performance planning, assessment, coaching, and development (Berkeley 2013, online).

Lastly, I would mention the concept of *administrative* control (or *work practice* controls), a general term indicating the *social aspect* of the control of agents as *individuals* who must operate in respect of the limits imposed by *responsibility* and *accountability*, which circumscribe the behavior of the agents and organs.

> Administrative controls develop control, since they direct employee behavior through the organizing of individuals and groups, the monitoring of behavior . . . and make employees accountable for their behavior . . . and the process of specifying how tasks or behaviors are to be performed or not performed (Green and Welsh 1988, p. 293).

> Procedures and policies, on the other hand, relate to the "bureaucratic approach to specifying the processes and behavior within an organization" (*ibidem*, p. 294).

The administrative control also takes on the significance of a control for the safety of the agents as workers in the organization, who must not suffer damage in carrying out their responsibilities. Finally, the term also has the bureaucratic meaning of *compliance control* regarding the operations and actions of the agents; this control is carried out by systems of recognition that verify that the acts are not contrary to the internal or the external rules of the organization.

Management control systems, however understood, are not always effective due to the fact that they control individuals in the organization who may not modify their behavior in the manner expected by the managers as a consequence of the action of the chosen control levers. Very often, no matter how well designed and structured a framework of management control systems in an organization may be, the results may not meet expectations, especially when some of the individuals subject to control band together (Hofstede 1978), in some cases bordering on a true rebellion, to sabotage the control actions. This situation has given rise to the expression "illusion of control":

> . . . the illusion fosters the belief among managers that conventional controls . . . accurately and validly measure, and thereby help determine, behavior. The illusion reflects a presumption that management can intervene when necessary and successfully effect change. Further, the illusion provides for the belief that, by changing a given mix of existing controls, managers make necessary and sufficient functional responses to internal or external change. To those managing with an illusion of control, negative consequences of managerial action often signify the necessity for more controls (Dermer and Lucas 1986, p. 471).

8.5   The *Rings* in *Macro* Management Control: The Objectives of the Strategic *Rings*: EVF and EVA

The MOEST model in Fig. 8.10 helps us understand that business profit organizations (firms) can be viewed not only as autopoietic and vital systems but also from two other points of view:

1. As *instrumental* systems through which people carry out the economic activities indispensable for production, consumption, savings, and investment of wealth

2. As *teleological* systems, since they are units with an autonomous existence with respect to the individuals operating within them; thanks to the managerial transformation they follow a trajectory in the production, economic, and financial space of which they are a part in order to achieve a system of vital objectives

The *ultimate aim* of organizations, which are viewed as *instrumental systems*, is the maximum achievement of the stakeholders' goals through a system of high-level objectives which I shall define as *institutional goals*. Satisfying the *institutional goals* represents, more than the *ultimate* objective of the organization, a true *condition of its existence*: the organization can continue to exist only if it achieves its *institutional goals*, thus leading to a long-lasting maintenance of the conditions of "unlimited life." In particular, the firm can continue to exist only if it produces "technical transformations" that can satisfy:

(a) The *clients*, through the production of prized goods or services (and thus possessing greater value with respect to the factors used to obtain them), of adequate quality, continually innovated, and at a price kept as low as possible with respect to competitors
(b) The *workers*, by offering them *monetary payments* that can lead to *cooperation* and *coordination* and a *work environment* that can best satisfy their professional aspirations and favor a *career* marked by commitment and loyalty toward the organization
(c) The *suppliers*, by guaranteeing stable supplies and punctuality in fulfilling orders
(d) The *suppliers of financial capital*, debt, by assuring a return in line with those offered by the market and, above all, security in the payment of interest and capital reimbursement
(e) The *suppliers of risk capital*, equity, by sufficiently increasing value through the production of high rates of return
(f) The entire *collectivity*, external stakeholders, by favoring employment, thereby guaranteeing a tax flow over time and safeguarding the environment by operating under conditions of sustainability

In order to achieve the *institutional goals* and respect the *constraints* set by the stakeholders, the TR4-M managerial transformation must provide the organization with specific control systems of various levels, which are *interconnected* to form a network and which often interfere with one another. The TR4-M must also define a system of coordinated *directional (managerial) objectives* that can be expressed quantitatively and represent a reference for control systems; the latter must ensure that the *directional objectives* are reached as well as maintain the desired *performance levels*, guaranteeing not only the *results* but also the operational *standards* of the processes and operational units. The *directional objectives* can be considered in all respects as *performance objectives* of the "technical" transformations. The system of control systems and directional objectives that must be achieved form the overall general management control system, which carries out two forms of control: *macro* and *micro* management control. These two aspects are distinguished
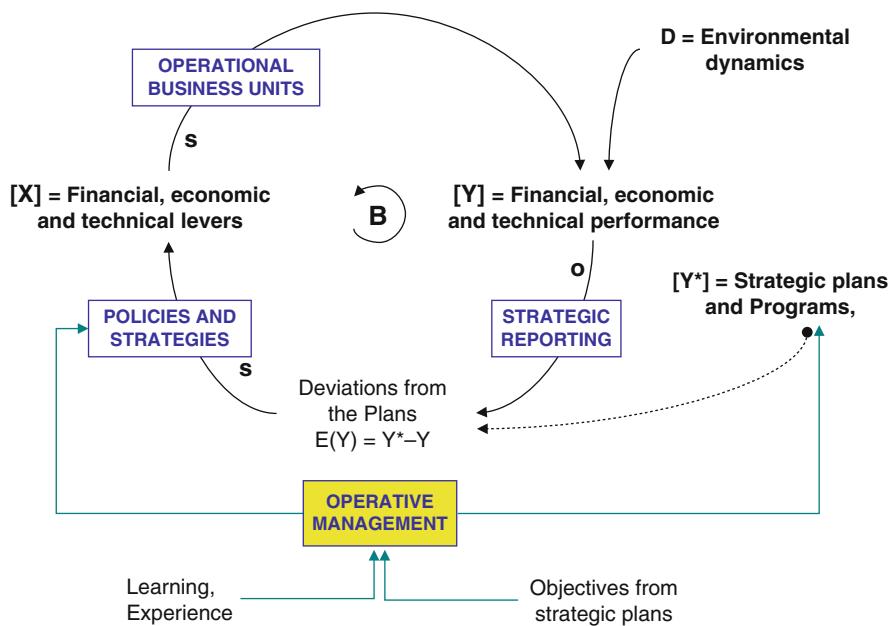
Fig. 8.13 General model of strategic entrepreneurial control of management in production organizations

not so much in a logical sense but in terms of the type of specific control system used, the objectives to achieve, and the time scale over which they act. *Macro* management control takes on the form of a *strategic control* as well as an *operational control*. We have already seen in Fig. 8.11 that the *strategic control*, after having received the *vital* objectives from the *entrepreneurial transformation*— which are conditioned by the policies and limits set by the CORPORATE GOVERNANCE and the stakeholders, in particular the stockholders—formulates the *strategic plans* and *programs*, which specify the levers and strategies to pursue the vital objectives and the top priority policies.

Figure 8.13 specifies that these plans, which cover a multi-year period, represent for the strategic control system the maximum objectives management must achieve through the technical, economic, and financial *levers*.

Figure 8.14, which is derived from Fig. 8.13, provides a concise representation of an *operational control* operating through the budget. In the practical application of the *budgetary control*, the model in Fig. 8.14 is to be understood as referring to a multitude of analytical control systems, each of which must act to achieve *specific values for the objectives*, which are specified for each *segment* of the budget. Only through analytical control systems is it possible, for each specific objective value (for example, $Y^* =$ Sales of "alfa" product in "northern" area in the month of "April" for the sales division "Hamburg"), to identify the values attained, Y, and the deviations, $E(Y) = Y^* - Y$, and to decide on the actions, $[X]$,
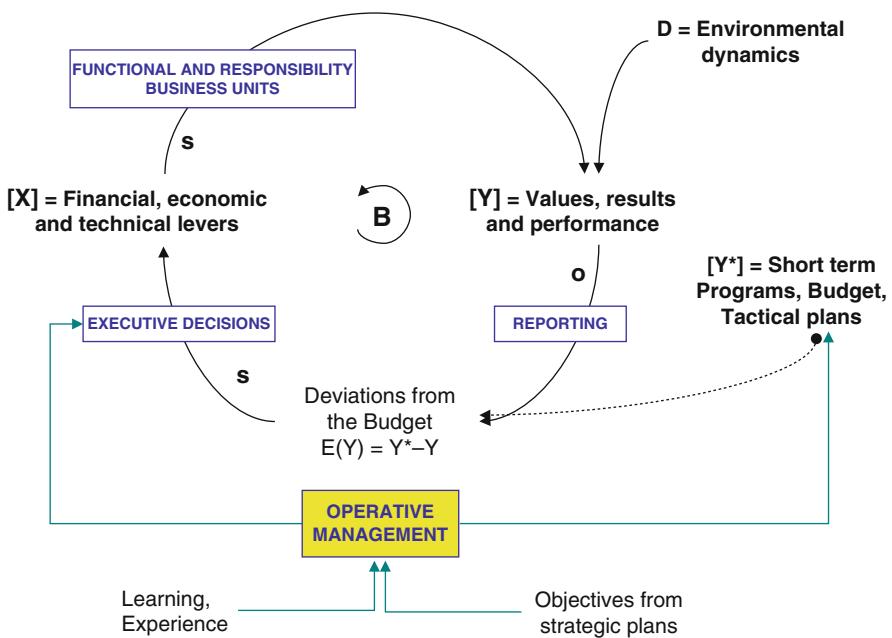
Fig. 8.14 General model of operational control of management in production organizations

in order to eliminate $E(Y)$ or accept it. In this sense the control through the budget, even more so than the strategic control, requires the activation of a true *system* of coordinated parallel control systems to achieve the general objectives regarding planned profitability. The *continual reporting* detects the trajectory of the values produced by the organs, the results, and the performance of the operational management; subsequently, through a comparison with the budget, it determines the deviations, $E(Y)$, from which management gains information to make corrective decisions regarding the activities of the operational units.

Though a detailed analysis of the system of strategic entrepreneurial *performance* objectives is not possible here, some further brief considerations are useful (for a more detailed discussion, see Mella 2005b, 2011a). In small- and medium-sized firms, where *equity* is concentrated in the hands of a few "owners," the highest order directional objective can be considered to be the achievement of the *maximum flow of profits* from the most limited use of *equity* possible. The *roe* becomes the main indicator of *economic performance*, as long as it is in line with the institutional goals and compatible with the *financial constraints* regarding the raising of the capital necessary for survival and organizational growth while also respecting the *monetary constraint* of a balanced cash management. For large corporations, where the capital is controlled by a large number of shareholders who can exchange their shares on the capital markets, the *entrepreneurial* transformation requires that the *managerial* transformation pursue *as* the *maximum* performance objective the *continual increase in*

shareholder value; this objective translates into the managerial objective of obtaining *profitability* levels that maximize the EVF, since the EVF defines the theoretical value of share trading.

In short, recalling that the value of an amount of non-redeemable capital, K, corresponds to the present value of its future average income, R, discounted at the rate "i" based on the well-known formula for discounting a perpetual income stream ($K = R/i$), then, in a totally analogous manner, the *EVF* can be calculated as follows:

$$\text{EVF} = P^*/ce^* \qquad (8.4)$$

where

– P^* indicates, for simplicity's sake, the expected future average profit of the firm before taxes (EBT), calculated with reliable and realistic multi-year plans.
– ce^* indicates the *cost of equity*, that is, the expected *financial interest* or *dividend yield* that would convince the shareholder to invest *equity* in the firm, a yield the company must guarantee in order to keep intact the shareholder's investment; thus, for the firm, ce^* theoretically represents the cost of attracting and maintaining equity.

Indicating by E the amount of the original *equity* investment which is assumed to produce P^*, then the ratio

$$roe^* = P^*/E \text{ (return on equity)} \qquad (8.5)$$

represents the highest order *managerial performance objective* of the firm, since it allows us to quantify EVF, as can be seen from the simple comparison of roe^* and ce^*. In fact, substituting P^* from (8.5) into (8.4) we get

$$\text{EVF} = E \,(roe^*/ce^*) \qquad (8.6)$$

If $roe^* = ce^*$, the *entrepreneurial* transformation guarantees the monetary integrity of the equity invested, since EVF = E; only if $roe^* > ce^*$ does the firm produce value for the investor, since from (8.6) it follows that EVF > E. The difference $G = \text{EVF} - E$ is defined as the synthetic value of *goodwill*, representing the *capital gain* to the capitalist if the firm is sold.

From (8.6) we immediately observe that the highest order *managerial performance objective* is to obtain a roe^* greater than ce^*; this managerial objective is, in turn, more precisely defined in the additional *managerial* objective of producing revenues, R^*, and incurring costs, C^*, that enable the firm to produce P^*, which quantifies the roe^* necessary to increase EVF.

The calculation of *cost of equity* is not simple, but it should generally correspond, for example, to the average *roe* of the firms operating in the *same sector*, under conditions of equal risk, if we assume that the capitalist wants to invest *equity* in production firms of the same type, or correspond to the *risk-free* yield from

government securities, if we assume the capitalist–entrepreneur is interested in the security of the investment more so than in the interest rate on the capital. However, the *risk-free* yields must be adjusted to take into account the different risks in investing in the firm (Mella 2005b).

Observing the MOEST (Fig. 8.10), we can consider the growth in EVF, or G, to be the maximum survival objective the *entrepreneurial* transformation sets for the *managerial* transformation in order to enable the latter to determine the *operational* objectives of R^*, C^*, and P^* for planning and budgeting. EVF (or G) can be viewed as the *global driver of the production of value*, since it serves as an objective to assess the overall performance levels of the four organizational transformations that succeed the entrepreneurial transformation.

Along with EVF, as a driver of the production of value, economic value added (EVA), or *residual income*, is also frequently considered. EVA was conceived of by Stern Stewart & Co., a global consulting firm, which launched it in 1989 (Stern et al. 1995). This measure shows in a simple manner the *residual economic result* obtained by subtracting the cost of debt and equity capital from NOPAT, that is, the *net operating profit after tax*:

$$\text{EVA}^{®} = \text{NOPAT} - (I + ce^*E) = (\text{EBIT} - \text{Tax}) - rod\,D - ce^*E \qquad (8.7)$$

The innovation here is the fact that EVA is calculated by subtracting from NOPAT as the cost of capital not only the interest on debt but also the expected return to shareholders' equity. However, NOPAT does not correspond simply to after-tax EBIT before deducting the cost of capital but, except for particular cases, is determined by adjusting EBIT using a certain number of *key adjustments* to the book values in the income statement to take into account the costs that must be capitalized or the capitalized costs that should have been included in the calculation of the economic result.

> These adjustments aim to 1) produce an EVA figure that is closer to cash flows, and therefore less subject to the distortions of accrual accounting; 2) remove the arbitrary distinction between investments in tangible assets, which are capitalised, and intangible assets, which tend to be written off as incurred; 3) prevent the amortisation, or write-off, of goodwill; 4) eliminate the use of successful efforts accounting; 5) bring off-balance sheet debt into the balance sheets; and 6) correct biases caused by accounting depreciation (Young 1999, p. 8).

Taking into account (KA), NOPAT is calculated as follows (P indicates net profit):

$$\text{NOPAT} = (\text{EBIT} - \text{Tax}) + \text{KA} = (P + I) + \text{KA} \qquad (8.8)$$

Ignoring KA, EVA then takes on the simpler form:

$$\text{EVA}^{®} = (roe\,E + rod\,D) - (rod\,D + ce\,E) = (roe - ce^*)\,E \qquad (8.9)$$

From (8.9) we clearly see that EVA > 0 only if (*roe* > *ce**), that is, if actual *return on equity* is greater than the *expected return for shareholders*.

There are also other entirely equivalent ways to formulate EVA, but (8.9) suffices to show that, in fact, the maximization of EVF implies the maximization of EVA and vice versa (ignoring the *key adjustments*). In my opinion, therefore, though it is useful in assessing from year to year the actual balance sheet results, EVA does not represent an *explicit* managerial objective, as demonstrated by the following relation (Mella 2005b):

$$\text{EVF} = \frac{R^* + \text{EVA}}{roe^*} = E + \frac{\text{EVA}}{roe^*} = E + G \qquad (8.10)$$

EVA's top-down approach works well at higher levels of an organization but becomes more difficult to implement below the strategic business unit level. ABC was designed to overcome this deficiency. To implement EVA at the lowest level of *a* firm's operations, capital cost may be traced to activities and then to objects, using the principles of ABC (Kee 1999, p. 4).

8.6   The Objectives of the Operational *Rings*: *roe** and *roi**

Despite the fact *roe* is an autonomous indicator of financial performance, for the purposes of *management control* it represents a *performance objective* derived from other *managerial objectives*, in particular *return on investment* (*roi*), *return on debt* (*rod*), and *debt/equity ratio* (*der*) (8.3), as clearly shown in Modigliani–Miller's (1958) well-known equation:

$$roe = [roi + (roi - rod)der](1 - \text{tax}) \qquad (8.11)$$

Equation (8.11) is translated in the model in Fig. 8.15, which makes it clear that the control system for *roe** is typically a multi-lever system whose levers can act on:

(a) *roi*, that is, raising the return on invested capital above *rod*.
(b) *rod*, to produce a *spread* = (*roi – rod*) which, if positive, can raise *roe* above *roi*, taking into account the financial leverage obtained from *der*.
(c) *der*, increasing this ratio until the *roe* objective is achieved, with *roi* given and assuming *rod* is known; in this case it is necessary to act directly on the efficiency of the TR3-F, financial transformation, in determining the appropriate proportion between debt and equity.
(d) Self-financing (Sfin = EBT – taxes – dividends) to achieve an adequate *der* and limit the financial burdens.
(e) The *corporate tax rate* (tax) to reduce taxes.
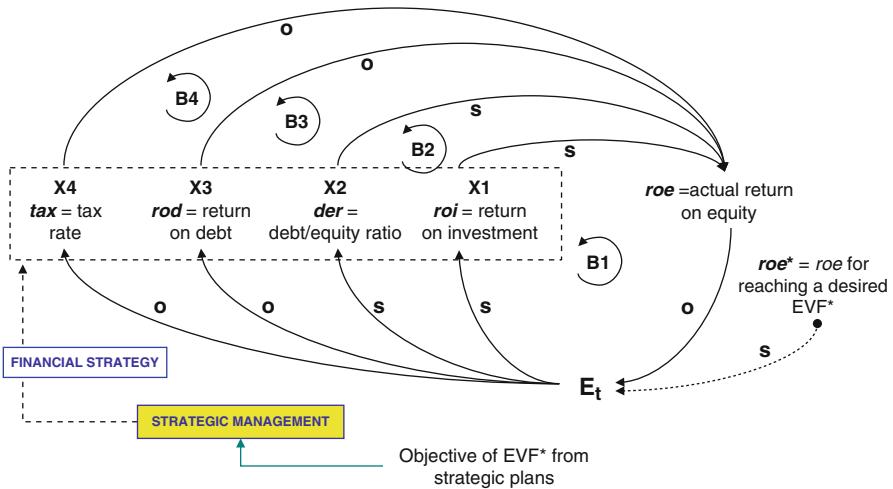
Fig. 8.15 General model of a control system for *roe*

The balancing loop (**B1**) in Fig. 8.15 clearly shows that the most important factor impacting *roe** is *roi*:

$$roi = \text{EBIT}/I = (R - C)/(D + E) \qquad (8.12)$$

which represents a new *performance objective* even more significant than *roe**, since it represents at the same time the *maximum economic and financial performance of the firm*. As a result, "*roi**" becomes the most significant *managerial objective* in terms of planning and budgeting, since the other levers are subordinated to it. "*rod*" quantifies the cost of debt capital, which is assumed to depend on the credit market; "*tax*" indicates the average income tax rate as determined by each country's fiscal policy. The choice of levers to control *roi* represents the medium- to long-term strategy which *management control* must specify and the *performance management* must carry out in order to achieve the necessary "*roi**."

While *roe** is controlled by the multi-lever control system shown in Fig. 8.15, the control system for achieving the objective for *roi** requires, as (8.12) demonstrates, the activation of two main levers which allow for the control of EBIT (numerator) and invested capital, $I = D + E$ (denominator), whose values, in turn, become the *managerial* objectives of the economic transformation. Of the two levers, EBIT is the main one; the optimal value of EBIT* that achieves the objective for *roi** is, in the final analysis, the *maximum strategic performance objective* of the *economic* and *productive* transformation, which must be controlled for the survival of the firm. However, (8.12) clearly implies that the control of EBIT must be carried out by a control system that acts through two important levers: *sales revenues* (SR) and *production costs* (PC), as summarized in Fig. 8.16.
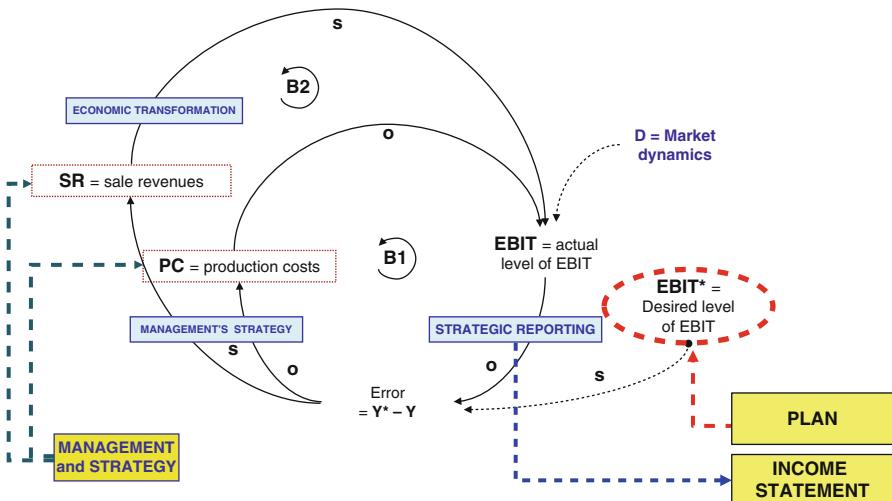
Fig. 8.16 General model of a control system for EBIT

The control system for EBIT carries out a *strategic management* control if it considers the EBIT flows for a multi-year future time period; it can be transformed into an *operational management* control if it covers only 1 year. In multi-business companies the control system for EBIT synthesizes a number of control systems which act in parallel fashion and all of whose objectives is the EBIT of each business.

A network of *lower level* control systems depends on these high-level control systems; the former operate in a coordinated and continual manner to control the *desired* levels of both *sales revenues*, SR*, and PC*, which ensure that the EBIT* objective will be met.

The control systems for achieving planned SR* and PC*, both globally and for each business, can activate various control levers in terms of product, market, price, technology, productivity, quality, distribution channel, etc. Therefore, the management for each control system must determine the appropriate activating *strategy*. Moreover, the various control systems for EBIT*, SR*, and PC* must achieve multiple, correlated (at times even often contrasting) objectives, so that it is necessary to establish not only a *strategy* but a control *policy* as well.

Figures 8.17 and 8.18 present in more detail the general models of the two control systems that operationally translate the revenue and cost levers indicated in Fig. 8.16. Such models can be specified for each business and each period of reference (year, quarter, etc.). In particular, Fig. 8.17 presents an example of a *multilayer* and *multi-level Ring* to control PC*.

The first level, operational control, is based on the two variables that can be impacted in the short run: the purchase price of factors of production and the unit factor requirements. Where there is a positive deviation $E(C)$ (greater-than-desired costs), management tries, on the one hand, to optimize supplies by reducing unit
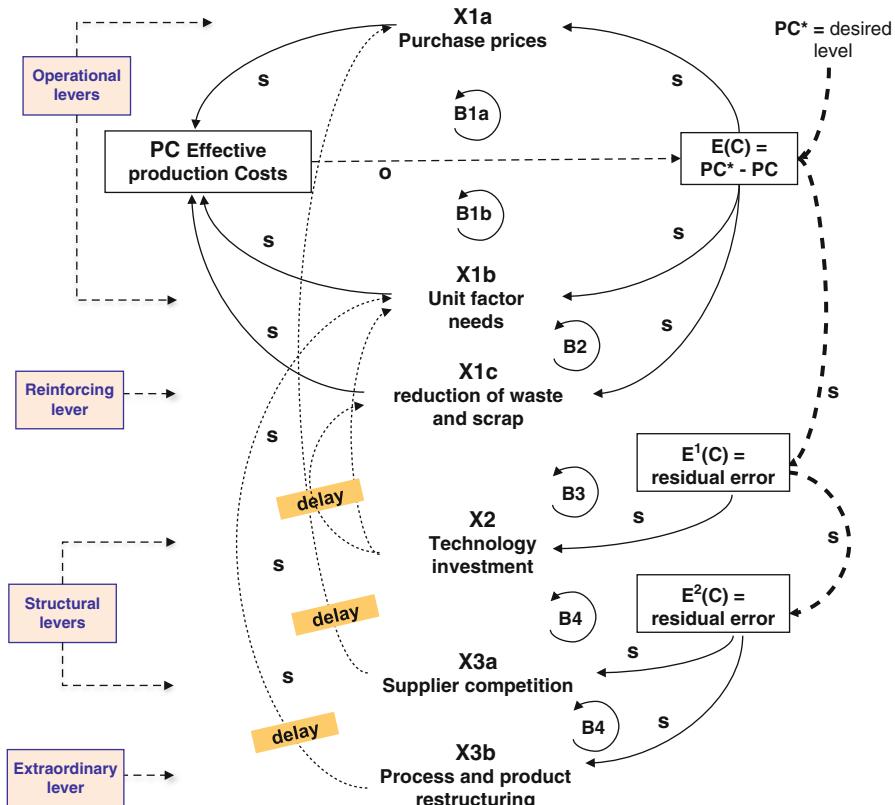
Fig. 8.17 General model of a multilayer system for controlling production cost

prices as much as possible and on the other to eliminate waste and redesign the product to reduce the unit requirements for material, labor, services, and other factors.

The other measures (the order in which they appear in the model in Fig. 8.17 is merely indicative) are more of long term and structural. The investment in technology and process restructuring levers is needed to save on the use of factors of production; the lever for changing the pool of suppliers, placing them in competition with one another or seeking out more favorable contract terms, is needed to impact the level of supply prices.

It is easy to imagine that these wide-ranging *structural* interventions most likely will not directly impact the *budget*; nevertheless, the model in Fig. 8.17 is so general as to be applicable also in cases of other forms of programming and control: *target costing* and *activity-based management*, and others still which I can only mention.

Figure 8.18 outlines the *multilayer* model of a *Ring* to achieve the desired level of revenues.
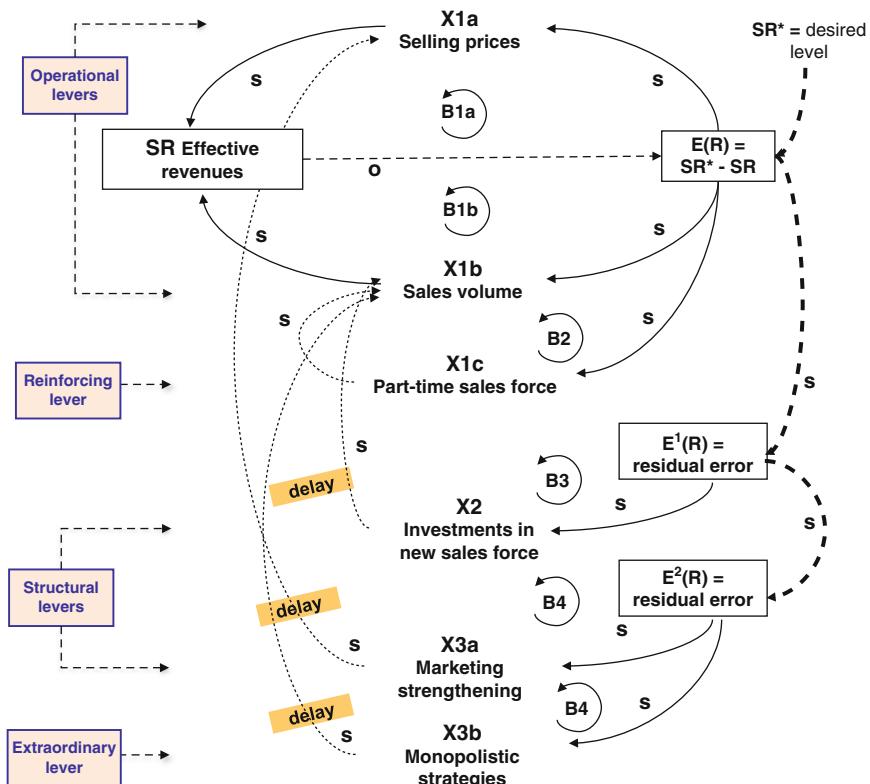
Fig. 8.18 General model of a multilayer system for controlling sales revenues

This case also calls for various levels of intervention: the operational control entails the price and sales volume levers. The other action variables permit a medium- to long-term *structural* control. In particular (the order is merely indicative), the strengthening of the sales force and determination of monopolistic strategies (or, in any event, strategies capable of increasing the advantage from size) affect sales volumes; the marketing action, which entails the strengthening and control of the respective variables (advertising, promotions, discounts, packaging, etc.), influences the selling prices.

Figure 8.19 combines the sales revenue and production cost *Rings*; it derives from Fig. 8.16, after substituting the synthetic PC and SR levers with the corresponding *Rings* in Fig. 8.17 and 8.18. Thus, the model in Fig. 8.19 represents the firm as a *multi-objective* control system in which management must try to simultaneously achieve the cost and revenue objectives, PC* and SR*, based on a well-delineated *policy* aimed at achieving the "goal" represented by the final *roe** and EVF* objectives. The *Ring* for the EBIT* in Fig. 8.19 is based on planning and programming, which is joined by the parallel operation of a process to *detect*
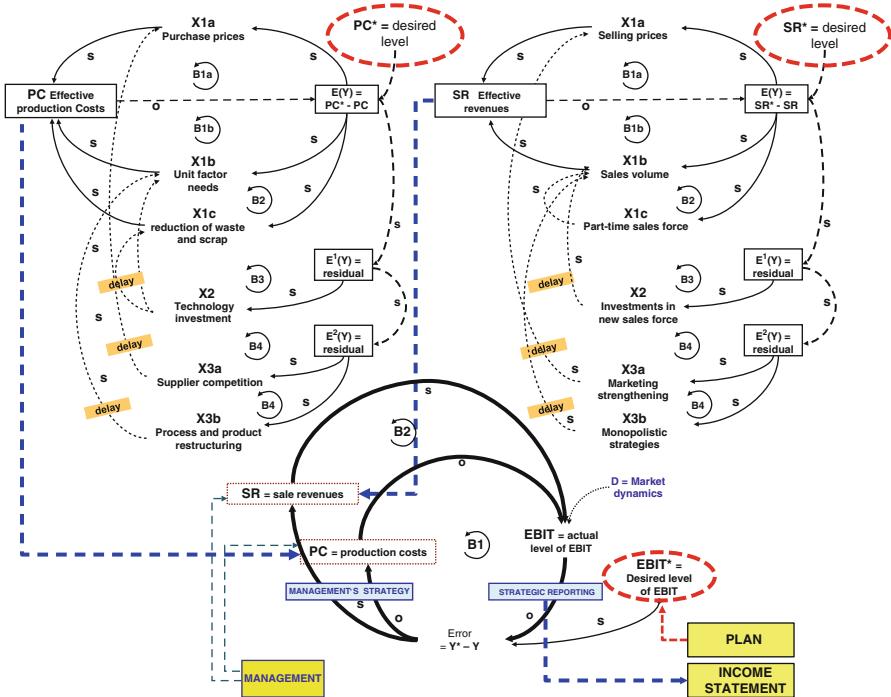
Fig. 8.19 General model of a multi-level, multi-objective control system for a firm's EBIT (see Figs. 8.16, 8.17, and 8.18)

performance and to evaluate it by identifying *variances* and deciding on the *strategic levers* for containing them.

We can easily observe the similarity between the structures in the models in Figs. 8.17 and 8.18. This is no coincidence; in fact, it is easy to note that the two *Rings* are entirely similar and follow the pattern already outlined for the control system of a car's speed in Fig. 4.15. Moreover, the similarity between the model in Fig. 8.19 and the control system for the flight of an airplane in Fig. 4.16 is entirely evident. Even the symmetry between the models in Figs. 8.17 and 8.18 should not surprise us; it is clear observing TR2-E in the MOEST (Fig. 8.10) that the control of revenue flows is (almost) symmetrical with respect to that of cost flows. We can intuitively deduce that the *operational control* of EBIT must achieve an even more analytical level by activating a network of increasingly more detailed *Rings* that operate in a coordinated and continual manner to control the level of the control levers in Figs. 8.17 and 8.18.

Figure 8.20 presents the control system for short-term *cash flow* (K), which originates with the cash inflow from *receipts* (KR variable) less the cash outflows from *payments* (KP variable).
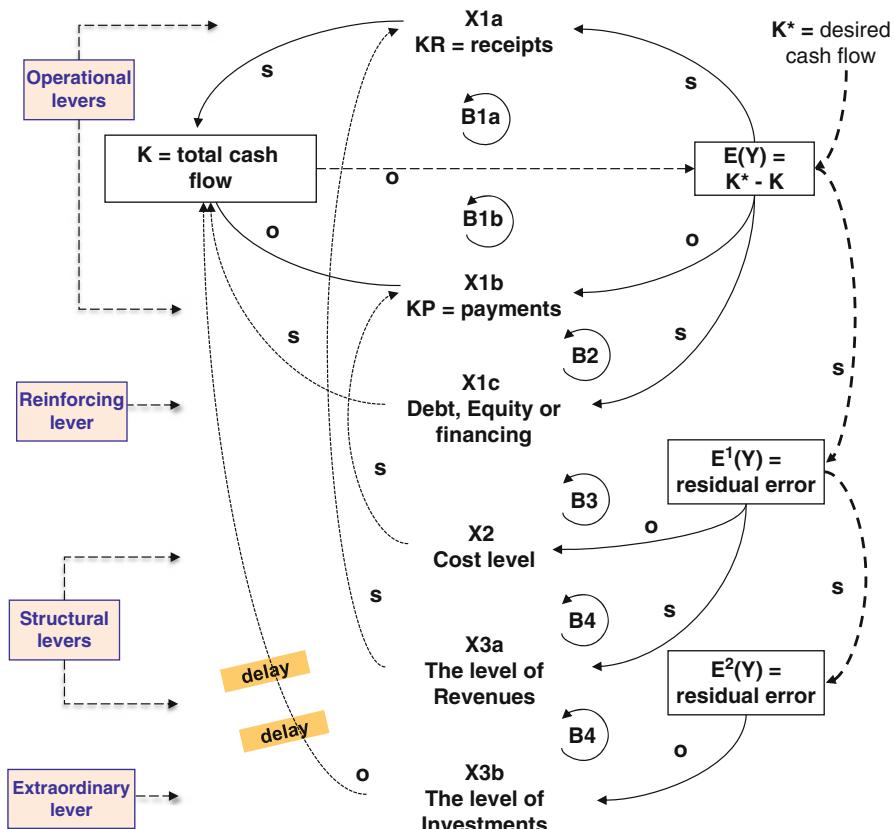
Fig. 8.20 General model of a control system for cash flow

The flows from revenues and payments represent the variables influenced by the current operational control of cash flow; if the restructuring of these flows were not sufficient to produce the desired volume of cash flow, then long-term and structural interventions would be called for that act on the other levers: investment, disinvestment, financing, etc.

In concluding this *section* I would remind the reader that TR5-I, the *entrepreneurial* transformation (Fig. 8.10), sets objectives not only for profitability but also for the firm's *growth*, understood as a long-lasting growth in sales revenue and invested capital. The growth of the firm requires that:

(a) Sales can be increased with a consequent absorption by the markets.
(b) Along with the growth in sales there is an adequate growth in the production processes and thus in the capital invested in the firm.
(c) The increase in the size of the production processes is adequately financed.
(d) The growth in sales and invested capital leads to a proportionate growth in operating income as well.

Organizational growth requires the availability of financial resources, which can be obtained from *outside sources*—loans or increases in capital stock—or internal ones, typically *operating cash flow* (or gross self-financing and net self-financing). The generation of *net self-financing flows* over time is possible only if the firm's revenue is sufficiently high to allow for a fair amount of dividend distributions while at the same time allowing part of the profits to be kept as a reserve and thus as self-financing. Once again the control of the performance of the *managerial* transformation can be further refined. It is not enough for *roe* > *roe**; it is also necessary for (*roe* – *roe**) $E \geq$ sfin*, where sfin* is the net self-financing needed to achieve the desired levels of growth.

For completeness sake, I would emphasize that the *managerial control* set up to maximize shareholder value and at the same time organizational growth is also referred to in modern management literature as value-based management, a new managerial approach aimed at the production of value for the shareholders (Mella and Pellicelli 2008).

Arnold's definition in this regard is significant:

> Value-based management is a managerial approach in which the primary purpose is long-term shareholder wealth maximization. The objective of a firm, its systems, strategy, processes, analytical techniques, performance measurements and culture have as their guiding objective shareholder wealth maximization (Arnold and Davies 2000, p. 9).

Copeland, Koller, and Murrin's definition is more specific:

> VBM is very different from 1960s-style planning systems. It is not a staff-driven exercise [. . .] Instead, it calls on managers to use value-based performance metrics for making better decisions. It entails managing the balance sheet as well as the income statement, and balancing long- and short-term perspectives (Copeland et al. 2000, p. 87).

Morin and Jarrel clearly refer to the double interpretation of VBM: a "mental attitude/selection of operational methods":

> [Value Based Management] is both a philosophy and a methodology for managing companies. As a philosophy, it focuses on the overriding objective of creating as much value as possible for the shareholders. . . . As a methodology, VBM provides an integrated framework for making strategic and operating decisions (Morin and Jarrel 2001, p. 28).

> Once the company develops strategies, a number of operational drivers that are key to implementing the strategy have to be identified. By focusing on these operational drivers, the company's strategy is successfully implemented, which in turn improves the value drivers, creating aggregate value (*ibidem*, p. 343).

The control systems presented in this *section*, specifically MOEST, are appropriate for summarizing the logic of value-based management in capitalist enterprises. According to the MOEST, we can highlight the basic strategy of the control system that value-based management must activate to maximize EVF and shareholder value:

1. Select those investments having a *roi* > *min roi**—that is, the level of *roi* sufficient to achieve *roe**—for the entire firm; if there is more than one, it chooses the one having the max *roi* and the *minimum payback period*.

2. Choose the investments that, in any event, have $roi \geq rod$ and $roi \geq wacc$ (weighted average cost of capital) and which are sufficient to guarantee min roe^*.
3. Choose financing with min $wacc$ and *min rod* (other conditions held constant).
4. If $rod < roi$, increase D and reduce E; turn to rule (1).
5. Substitute, when possible, investment I with J if $roi(J) > roi(I)$; in this way the average *roi* for the entire firm will increase.
6. Substitute, when possible, financing F with G if $rod(G) < rod(F)$ in order to reduce the average *rod* for the entire firm.

Trying to achieve the objective of increasing EVF means setting roe^* as the objective and using the levers shown in Fig. 8.15, which, in turn, represent lower level objectives whose attainment implies the activation of specific management *Rings*. In particular, roi^* becomes an objective for control systems for EBIT, which are presented in concise form in Fig. 8.19. Since net profit and EBIT derive from the "technical" transformations (productive, economic, and financial) indicated by the MOEST (Fig. 8.10), clearly it is necessary to consider a hierarchy of operational objectives for these three transformations, which will then be further detailed through operational goals to be achieved by the various business divisions in carrying out their operational processes. In particular:

(a) The *sales process* will be directed at increasing market share, price levels, percentage of market penetration, and so on.
(b) The *production* process will be aimed at objectives that specify the production and stock volumes, production cost levels, labor productivity, average production capacity employed, degree of flexibility of facilities and machines, degree of elasticity of machinery and equipment, replacement times, and so on.
(c) The *process for* purchases will be directed at specifying the volumes and rates of supplies, the minimum number of suppliers, the degree of independence toward certain markets, and so on.
(d) The *financing process* will aim at achieving objectives that determine the sources and cost of financing, the investment recovery period, and so on.

Thus, a "pyramid" of corporate objectives is formed, which is summarized in Fig. 8.21. In order to achieve these objectives the entrepreneurial transformation must draw up a *strategy* which the managerial transformation translates into annual *budgets* to be met by the *micro management control*.

8.7   The *Rings* in Micro Management Control: Stock, Productivity, and Quality

Micro management control, or *operational control,* concerns the various functions (supply, production, sales, logistics, personnel, finance, etc.) and various operational units (plants, divisions, departments, individual facilities and/or workers, etc.) which, by activating various levels of operational levers, produce the values

maximum EVF and maximum ROE

or, more realistically, a minimum level of profitability, ROE*, and a desired growth rate

objectives for ROE, ROI, ROD, DER, growth rate, afin

specified as

objectives for REVENUE (QP and Pp), VALUE ADDED, COST, MARKET SHARE, etc.

further broken down into objectives for

PURCHASES	PRODUCTION	FINANCE	SALES
optimize purchases and stock EOQ logistical optimization	increased return from reducing factors of production and production costs increased use of production capacity flexibility of machinery productivity and reduction of waste, etc. optimal production lot, etc.	reduction of financing requirements Debts Mix reduction of ROD increase of CIR optimization of DER	price increase/reduction increased market share and customer loyalty increased sales and customer satisfaction customer diversification, etc.

Fig. 8.21 The "pyramid" of corporate objectives

that must be achieved by those values chosen as objectives in the budget. The micro control concerns the achievement of short-term objectives (producing at an average cost below a certain level; fulfilling an order at a given cost and within a given time period; not exceeding the maximum or the minimum warehouse stock; reducing the average cost of capital for the next quarter, etc.) and makes use of various and well-structured instruments to detect the *variances* and determine which control *levers* to adopt. Here I can only mention a few of the many such instruments: annual and interim functional budgets and reports, time-based programming (PERT, CPM, Gantt, etc.) for projects, activity-based costing and activity-based management (Cooper 1989; Cooper and Kaplan 1991; Hubbell 1996), target costing, detection of times and methods, and so on.

Although we are not able here to examine the many *Rings* that make operational control possible, it is nevertheless useful to take a brief look at some of these systems which are particularly significant and which carry out a *micro management control* of the *productive transformation*: *stocks, productivity,* and *quality*, which are "technical" variables that require detection instruments and control levers of a "technical" nature.

Figure 8.22 illustrates the general model of the control system for *stocks* of a component or a material that is purchased externally; this is typically a system operating through impulses (Fig. 4.7). An example of a calculation of EOQ is undertaken in Sect. 8.9.4.

The control of the efficiency in the use of productive factors is fundamental to the efficiency of TR1-P, the productive transformation (Fig. 8.10). In order to detect the levels of *productive efficiency*, a number of indicators of the efficiency of
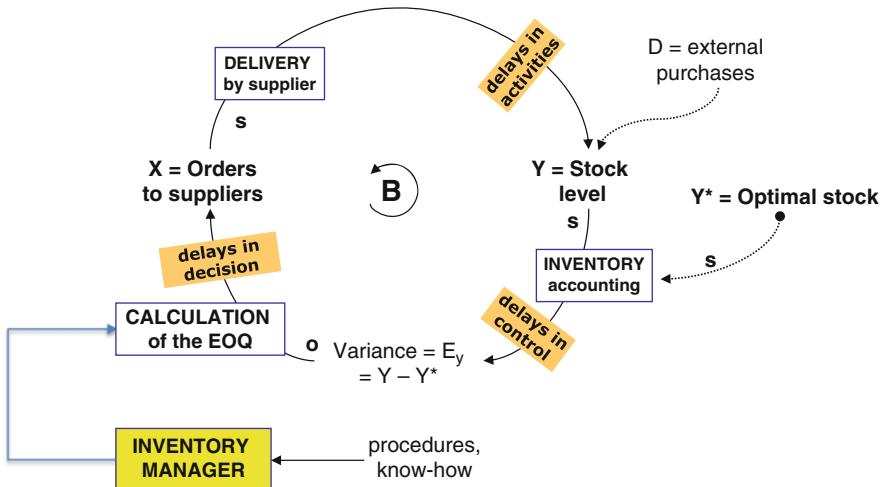
Fig. 8.22 Simplified model of a control system for the stock of a component

the employment of the factors in production have been created. Value analysis, laboratory tests, time-based programming of processes, time and method analysis, efficient scheduling of tooling, and maintenance procedures all enable the firm to control the efficient use of material and machinery.

Of particular importance is the control of labor productivity (production performance), for which purpose tools for the organizational analysis of responsibilities and time management techniques can be adopted along with various control levers, among which:

(a) An improvement in labor quality through selection, training, and retraining processes
(b) The rationalization of labor use through the reengineering of processes, flexibility in labor use, and labor mobility within the firm
(c) The control of worker performance by providing adequate motivation through monetary and non-monetary incentives (career, security, in-house day care, etc.) and an improvement in work conditions
(d) An improvement in quality regarding the use of materials, machines, and equipment

The control of *product quality* and the quality of *production and sales processes* are also fundamental to the *productive* transformation, whether the firm produces goods or services or produces by job order or process. The techniques for evaluating the quality levels and the *levers* to be used to control this variable are numerous and complex. Some brief comments in this regard can aid the reader who wishes to look further into this topic.

It is not easy to define the quality of a firm's production; three concepts are particularly significant.

1. *Functional (or market) quality* is the basis for *customer satisfaction* (Wellemin 1990); it defines the set of characteristics which, from the customer's point of view, make the *product* appropriate for *use*, that is, capable of satisfying a specific *use or utility function* of the good or the service (Band 1991), taking into account a desired standard of *reliability*—the product must provide use that is not interrupted due to imperfections—and *safety*—the product must be capable of use without damage or risks (Ebeling 1997).
2. *Design* (or *intrinsic, productive*) *quality* is the set of characteristics that, from an internal point of view (in terms of production processes), make all the product units conform to a standard of reference (prototype, sample, model, design), as stated by Crosby (1979), who defined quality as "*conformance to requirements*."
3. *Environmental* (or *context*) quality is the set of characteristics which, from the point of view of external impact, make the product compatible with the environment, in terms of pollution, waste disposal, environmental risks, or suitability for introduction into the context in question. Today, in our highly interconnected society, no product can avoid an *assessment of adequacy*: some cars are too polluting; some homes scar the landscape; some packages are too cumbersome; some scooters are too noisy; some motorcycles are too dangerous; some railway lines are too invasive, and so on.

Quality must be ensured for the customer as a condition of survival for the firm, since it influences considerably the efficiency level of the TR2-E, economic transformation (Fig. 8.10), both in terms of revenues, by influencing the selling price and demand level, and costs, since a variation in quality levels causes a variation in the cost of quality certification, protection, and restoration. All the corporate functions that guarantee *quality assurance*, on which *customer satisfaction* depends, contribute to the achievement of high levels of quality (Krishnaiah and Rao 1988).

The *levers* to control *functional quality* must certify that the product maintains (or increases) over time the overall *use function* required by the market. This type of control is carried out in the context of the marketing function, which must be capable of:

(a) Verifying and influencing consumer tastes (market research, consumer panels, pilot markets, etc.)
(b) Creating new needs and aspirations (advertising)
(c) Updating the product's technical features as the types of needs change (product design quality)
(d) Coming up with new features that can visibly differentiate the product so as to create a product "image" to serve as an instrument for satisfying consumer aspirations (brand, trademark, imitation effect, etc.)

Among the instruments used to control functional quality, *value analysis* deserves mention. This instrument starts from the principle that the value of a product will be interpreted in different ways by different customers, in terms of a high level of performance, capability, emotional appeal, style, etc., relative to its cost. This can also be expressed as maximizing the function:

$$\text{Value} = \text{Functional quality}/\text{cost}$$

Value analysis consists of a series of techniques and studies aimed at finding the best product composition in terms of the form and structure of the components and the necessary manufacturing processes, with the aim of minimizing the drawbacks resulting from the product's use (Crow 2002).

> Value analysis is a problem-solving system implemented by the use of a specific set of techniques, a body of knowledge, and a group of learned skills. It is an organized creative approach that has for its purpose the efficient identification of unnecessary cost, i.e., cost that provides neither quality nor use nor life nor appearance nor customer features. When applied to products, this approach assists in the orderly utilization of better approaches, alternative materials, newer processes, and abilities of specialized suppliers. It focuses engineering, manufacturing, and purchasing attention on one objective-equivalent performance for lower cost (Miles 1989, p. 3).

While the analysis of value is a technique that encompasses the design and manufacturing phases, *quality circles* (Ishikawa 1980) are organizational instruments for the control and improvement of quality as part of the quality function deployment. Gupta (2001, p. 771) states that a quality circle is:

> [...] a small group of people doing similar work who meet voluntarily and regularly, usually under the leadership of their supervisors; they identify and discuss their work problems.

Who better than the workers involved in the production process to identify the problems in assembling, stocking, and distributing materials, component parts, subassemblies, and other items at any stage in the manufacturing process? Who better than these workers to suggest solutions to eliminate such snags? While the *quality circles* rely on the *opinions* of persons *inside* the firm, the *panel of consumers* technique directly seeks out the opinions of *customers*, forming a panel of consumers who are chosen based on statistical sampling techniques and then asked to try out the product—usually a new one—so that they can use it and express an opinion on its quality (and, in general, also on the price/quality relationship).

The *levers* to control *design quality* must maintain a uniform technical standard in space and over time, which implies preventing or eliminating defects and setting up a service to carry out inspections and make substitutions, revisions, and repairs. Such *levers* translate into actions aimed at:

(a) Controlling the quality of materials and components, which involves the supply function.
(b) Controlling manufacturing; this control is carried out at the production function level.
(c) Controlling and testing the finished product (task of the production manager); in firms with product managers, the responsibility for controlling the finished product falls to the product head, though the control must be extended to the product process and its personnel.
(d) Controlling commercial distribution in terms of assessing promptness and precision in deliveries and the respect of sales procedures.
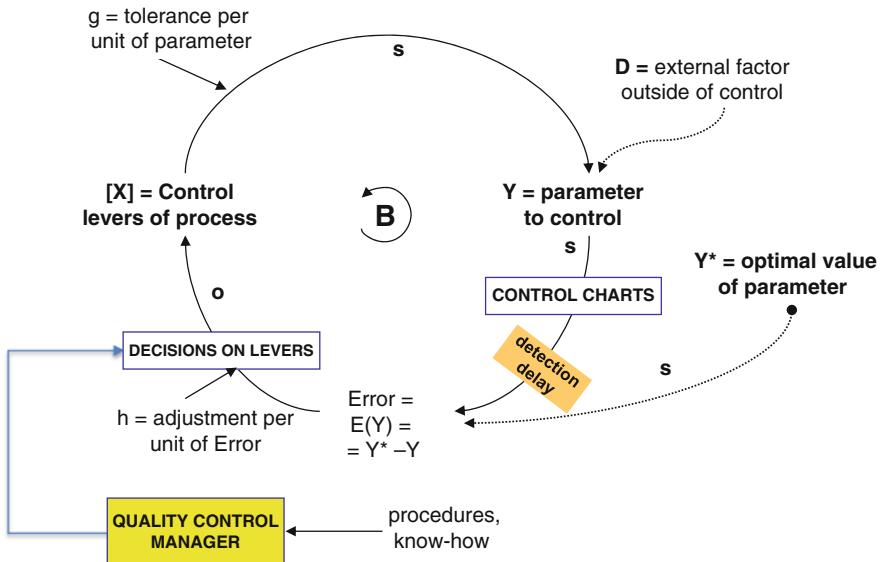
Fig. 8.23 General model of the control system of a parameter

The control systems for *design quality* verify, in particular, that the production processes maintain defects within the tolerance limit objectives (considered to be optimal); the systems undertake a sample survey regarding the inflow of materials and the outflow of production, as shown in the simplified model in Fig. 8.23. This survey uses *control charts*, which are special charts used when it is necessary to control a continual process. The control charts indicate the checks carried out, the measures obtained, and the upper and lower limits of the maximum tolerance measures for a given characteristic of the unit or for the production obtained from the process. The control charts allow for a simple and immediate diagnosis of the production process problems, on the condition that these problems are obtained *on line* (while the cycle is in progress) and in real time, that is, before the last moment has passed to intervene efficiently in the process.

In concluding this *section*, a brief mention is in order regarding the control systems that use structural levers through organizational restructuring and the *reengineering* of processes, organs, and functions. These are long-term levers that lead to radical changes in the organization but which can increase the power of control systems for processes (Hammer and Champy 1993; Davenport and Stoddard 1994). Equally potent, but more gradual in moving toward the desired objectives, is the *continuous improvement* which leads to the gradual improvement in the performance of processes leading to the objectives (which normally involve quality) (Deming 1982).

8.8   The Role of Planning in Transforming the Organization into a Cognitive, Intelligent, and Explorative System

Strategic entrepreneurial control (multi-year plans and programs) and *operational managerial control* (operational programs and budgets) have the same function and aim: to determine a specific *direction* for the firm in the *productive, economic, and financial space*. In fact, since they specify in detailed manner the operations selected to be carried out to achieve the values of the final objectives (revenues, supply costs, investment, cash flow, etc.), the multi-year plans, programs, and budgets (though at different levels of analysis) are the formal representation of the *programmed direction (trajectory) of the firm system in productive, economic, and financial space*, defined for the various sub-periods of reference.

However, these documents represent not only the fundamental instrument for the system's *direction* but also the indispensable condition for the *management control*, which seeks to maintain the trajectory in the established *direction* through specific control systems. In particular, only a well-structured budget with continual and timely revision along with multi-year plans will allow a control system for management to exist. For an outside observer, such a system turns the firm into an *automatic control system* whose *management*, represented by the entrepreneurial and managerial transformation (Fig. 8.10), is an integral part that "learns from its own errors" in a continual and mutual interaction with the external environment. Reporting plays a fundamental role, as it allows the firm to detect the variances in the values produced with respect to the budgeted values and to accurately analyze these to determine the causes of such variance and decide on the control levers to activate. Just as the *budget* must be precise and continually revised, the *reporting* must be precise and timely; it would be of no use knowing February's sales data in June or the production data after 3 months.

Feedback control, which is typically *reactive* and based on eliminating "system errors" from the programmed path, cannot, and must not, be the only type of control. *Cognitive transformations* (entrepreneurial and managerial, Fig. 8.10), through an accurate control to forecast the fundamental variables, must achieve a *feedforward control* by taking *proactive* actions to *prevent* variances, accurately deciding on the input levels that will produce the programmed output. In any event, *feedforward* control assumes a planning process which is continually updated and controlled through *feedback*. Planning and the budget are the *result* of the logic of *feedforward* control and represent the *instrument* for *feedback* control.

Planning in general, and the *budget* in particular, have other relevant functions, first and foremost as *instruments of cooperation and coordination* among the members, organs, functions, and processes of the organization, as foreseen by Beer's VSM (SS2) (Fig. 8.8) and by the MISTE (TR4-M) (Fig. 8.10). The budget is a tool for *coordination* since it is composed of functional budgets, which are divided into analytical budgets, all of which are coordinated by the economic calculation to produce a coherent set of values capable of verifying whether or

not the objectives have been achieved. As a result it is also an instrument of *cooperation*, since the achievement of the individual budgets by those in charge of them automatically entails the achievement of the final objectives of the firm.

The budget is therefore a basic performance management instrument since it allows the firm to control that the personnel and organs in the organization actually carry out their *responsibilities* and *tasks* in a coordinated and cooperative manner in order to achieve the various sections of the budget. However, for performance control the firm must be divided into *responsibility centers*, each of which with its own objectives to achieve, and include a system of incentives to ensure that personnel follow established levels of organizational efficiency. This form of control of responsibilities is also known as *task control*, which Robert Anthony defines as follows: "*Task control is the process of ensuring that specific tasks are carried out effectively and efficiently*" (Anthony 1988, p.12).

To restore cooperation and coordination various control levels can be activated: from participating in decisions to a rewards/punishment system; from creating an efficient information system to improving working conditions; and from worker participation in the firm's capital to forms of identification of individuals in the mission of the organization they belong to.

In order to carry out the management, coordination, cooperation, and organizational control functions the plans, programs and budgets must not be the result of a simple bureaucratic process directed from the top (cognitive transformation) down (technical transformations) but rather of a democratic participatory process, from bottom up, so that decisions are accepted by the entire organization, since they are formalized through the consensual participation of the operational organs.

Faced with environmental change—whose visible features are the increasing speed of the production of "innovation"; the increasingly larger geographical areas over which the various organizations interact; and the more crowded network of inter-organizational and intraorganizational relations that make it more difficult to interpret the overall behavior of each system—the capacity to achieve *teleonomic behavior* to ensure long-term survival depends on the ability of organizations and managers to understand change and to regenerate the cognitive processes, thereby producing a long-lasting autopoietic behavior.

If the firm, through the entrepreneurial and managerial control systems, learns from its errors, then the *plans*, *programs*, *and budgets become powerful tools of organizational learning*, since they will be based on:

- A careful analysis and evaluation of stakeholder expectations
- A process leading to a deep awareness of the organizational structure
- An accurate search for information, generally concerning future developments and forecasts regarding the environmental context, operating structure, and contingent situations
- A precise recognition of the external environmental constraints and the internal organizational ones
- The specification of the fundamental variables of interaction with the environment

As a result, the *planning* and *budgeting processes* move the firm to specify a system of dynamic objectives regarding productivity, economic efficiency, and profitability that guide management operations based on a typical *pull logic*; moreover, based on an equally typical *push logic* these processes help to gradually improve performance to permit the achievement of those objectives through performance levels continually adapted to a changing environment.

> The repertoire of tools that top managers use... is determined by their experiences. A dominant logic can be seen as resulting from the reinforcement that results from doing the right things ... they are positively reinforced by economic success (Prahalad and Bettis 1986, pp. 490–491).

In this sense the activation of strategic and operational control systems represents the fundamental moment of reflection of the entire organizational activity: a continual critical rethinking of the mission, competitive position, and decisions to be made. Thanks to participation and sharing in discussing the underlying hypotheses of the plans and programs that set out the control objectives and create the conditions for coordination and cooperation, the entire organization becomes involved in a deep, collective learning process regarding the internal processes, constraints, limits, potentialities, and functional relations with the other processes and with the environment. If carried out with broad-based participation and motivation, this process will play a fundamental role regarding knowledge and result in the improvement at all levels of the organizational structure, leading the firm to behave according to the logic of *learning organizations*, which was briefly discussed in Chap. 1. From an instrument of *modeling to predict and control*, planning becomes a process of *modeling to learn* (de Geus 1988, 1992):

> Institutional learning is much more difficult than individual learning. The high level of thinking among individual managers in most companies is admirable. And yet, the level of thinking that goes on in the management teams of most companies is considerably below the individual managers' capacities. [...] Because high-level, effective, and continuous institutional learning and ensuing corporate change are the prerequisites for corporate success [...] the critical question becomes, "Can we accelerate institutional learning? (Arie de Geus 1988, p. 70).

> The only relevant learning in a company is the learning done by those people who have the power to act (at Shell, the operating company management teams). So the real purpose of effective planning is not to make plans but to change the microcosm, the mental models that these decision makers carry in their heads. And this is what we at Shell and others elsewhere try to do (*ibidem*, p. 71).

> Our exploration into this area is not a luxury. We understand that the only competitive advantage the company of the future will have is its managers' ability to learn faster than their competitors. So the companies that succeed will be those that continually nudge their managers towards revising their views of the world. The challenges for the planner are considerable. So are the rewards (*ibidem*, p. 74).

As a system of efficient transformation (MOEST) the organization becomes in all respects a *cognitive, intelligent,* and *rational* economic *agent* with the capacity to *control* its own structure, processes, and dynamics in order to achieve ever-higher levels of efficiency. It is an *economic agent* because the organization–enterprise

designs and follows a unique trajectory in the *productive, economic,* and *financial* space in which it operates, even though it is characterized by a multitude of organs, processes, and businesses, as VBM has shown.

The organization is an *intelligent cognitive agent* in that, as we have seen (SS4 and SS5 of the VSM, and TR5-I of the MISTE), the organization engages in a *cognitive* activity aimed at providing meaning to environmental stimuli, translating these into information and, through programming, structuring this into knowledge, adapting to the changing environment while maintaining its identity through a continual autopoietic process, as prescribed by the autopoietic approach. It is a *rational agent* in the sense that *cognitive* activity must lead to maximizing the *efficiency* of the vital transformations by searching for the best productive, economic, and financial performances (Fig. 8.10).

However, we can also consider that the organization–enterprise, being a *rational, cognitive agent*, is also an *explorative agent* (Kauffman and Levin 1987; March 1991; Lewin et al. 1999; Mella 2006) that, in continually searching to improve performance in any way possible (*rational* agent), *explores* its territorial environment which may be divided into *areas of interest* (continents, states, regions, provinces, etc.) and "moves" toward those *areas of greater attractiveness* (Scott and Bruce 1994; Gephart et al. 1996).

Let us assume that the organization–enterprise, based on its knowledge (*cognitive, intelligent* agent), can divide the "territory that can be reached through its processes" into significant areas and that for each of these it can determine the value of the performance indicators it wishes to maximize. In this way the "territory that can be reached" (for each area, and subarea, in which it is divided) is characterized by a *function of attractiveness* which indicates the average level of the performance indicators held to be significant, thereby forming an *attractiveness landscape* indicating which areas are most attractive and which less so in relation to the various performance indicators that were chosen. Based on the characteristics of each area, it is possible that the *attractiveness landscape* may reveal "valleys" of moderate attractiveness, "peaks" of high attractiveness, or undesirable "pits" (no attractiveness whatsoever) to be avoided.

Therefore, continuing to assume that we wish to choose *roe* and *roi* and their components (Sect. 8.6) as performance indicators, it is likely that an area rich in potential consumers and with few competitors will be highly attractive due to potentially high revenues, from the side of both quantity and price, and thus offer a high *roi*. On the other hand, an area with many competitors would not be very attractive since, because of competition in terms of price and quality, we can assume a lower *roi*. An area with a low tax inflow would, however, have a higher *roi* compared to the one with higher inflows, with all other factors held constant. An area with heavy pedestrian traffic could favor sales for a small retail firm, while the one with ample parking facilities could increase the economic and financial performance of a large retail trader. Note that, to an *explorative* agent, the areas of attractiveness should not be solely conceived of as areas in a geographical territory but also as areas in the economic-financial environment, which is made up of businesses,

Areas of a chosen territory: (Ai), i = 1, 2...	Index of attractiveness (aij),	Performance indicators (Pi) evaluated for each area based on the index of attractiveness	indicators of performance risk, (ri) for each area	Attractiveness landscape
[1]	[2]	[3]	[4]	[5]
A1	[f11, f12, f13, ...]	P1 = mix of [f11, f12, f13, ...]	r1	AL1 = P1 r1
A2	[f21, f22, f23, ...]	P2 = mix of [f21, f22, f23, ...]	r2	AL2 = P2 r2
A3	[f31, f32, f33, ...]	P3 = mix of [f31, f32, f33, ...]	r3	AL3 = P3 r3
A4	[f41, f42, f43, ...]	P4 = mix of [f41, f42, f43, ...]	r4	AL4 = P4 r4
...	...	...	...	...

Fig. 8.24 Process to determine the attractiveness landscape

markets, distribution channels, technology, and legislation. The assumption that the organization–enterprise is a *rational* agent leads to the following *optimal strategy*:

1. Explore all accessible territories and areas that can be reached by its productive, economic, and financial processes.
2. Determine the most important factors *of attractiveness* or *undesirability*, and evaluate these in economic terms by forming the *fitness landscape*.
3. Translate the fitness landscape into the *attractiveness landscape*.
4. Shape, update, and continually explore, looking toward the future, fitness, and at*tractiveness landscapes*.
5. Choose the area(s) with the greatest *attractiveness*, that is, with favorable conditions for an increase in *roe* and *roi* (for example, the ease with which new businesses can arise, greater sales volumes, expectations for better prices and supply costs, greater productivity and public subsidies, high levels of social protection, stimulating environment, abundance of infrastructures, lower tax burden).
6. Avoid the "pits," and attain the highest "peaks" of attractiveness, that is, those areas where the best performance can be achieved.

The formation process for the *fitness landscape* and the *attractiveness landscape* can be summarized in the table in Fig. 8.24.

Column [1] indicates the different *areas*, Ai, considered interesting by the firm or by one of its businesses. Column [2] includes the *factors of attractiveness* that characterize each area in relation to the type of firm or business. The vector $[a_{11}, a_{12}, a_{13}, \ldots]$ represents the fitness landscape of area Ai. Column [3] determines the value of the chosen *performance indicator*, calculated as the mix of values from column [2] carried out by the entrepreneurial transformation after careful evaluation. Column [4] specifies, if possible, a *value of risk* for each area (certain if $r = 1$; impossible if $r = 0$). The *index of attractiveness* for each area is determined taking into account P and r, contemporaneously; the simplest way is by multiplying the two indicators. The vector $[P_i \times r_i]$, $i = 1, 2, \ldots$, indicated in column [5], represents the *attractiveness landscape*, Al$_i$, for the areas in the chosen territory.

Once again, even when viewed as an *explorative* agent, the organization–enterprise represents a control system, since the objective of achieving a given performance forces the organization–enterprise to continue exploring until it identifies the areas
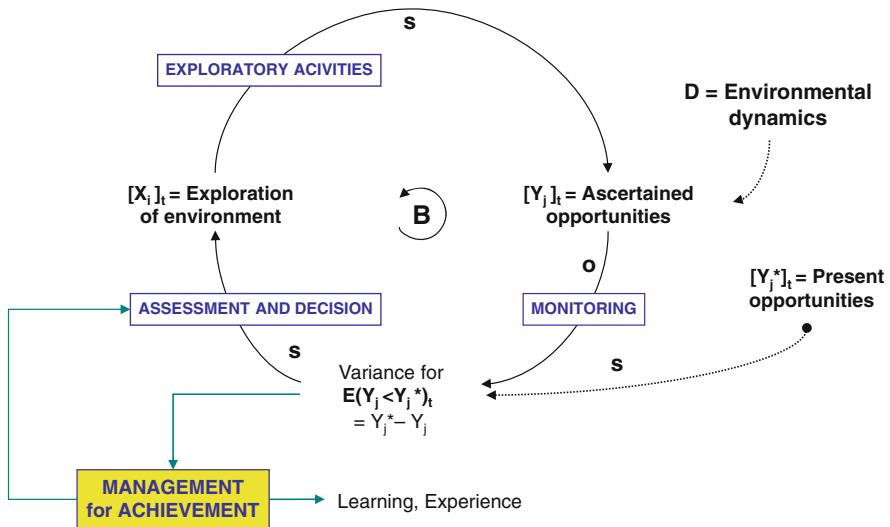
Fig. 8.25   Model of a firm as an exploratory system of achievement

in which the objective is attained. The control system can act in two ways: the firm *explores* the environment to *achieve* a given performance level, or it undertakes exploratory behavior to *improve* the previous performance levels. In the first case (achieving a given performance level) the control system can be defined as one of *achievement* and in the second (improving previous performances) as one of *improvement*. Systems of *achievement* can be represented by the *Ring* in Fig. 8.25, which is adapted from Fig. 8.13 or Fig. 8.14.

Systems of *improvement* are typical systems of pursuit whose new objective for time "$t + 1$" is the best performance achieved at time "t," to which management adds a plausible increment, taking into account competitors and limits imposed by the stakeholders, as shown in Fig. 8.26.

8.9   Complementary Material

8.9.1   *Dashboards in Performance Control Systems*

In formulating the VSM (Sect. 8.2), Stafford Beer observed that at all levels of the organizational hierarchy managers serve as filters of the information from the lower levels, since they synthesize this information in order to transmit it to the higher hierarchical levels, which in turn filter it again by sending synthesized information to even higher hierarchical levels. In a complementary manner, the further one moves toward the operational base of the organization the more analytical
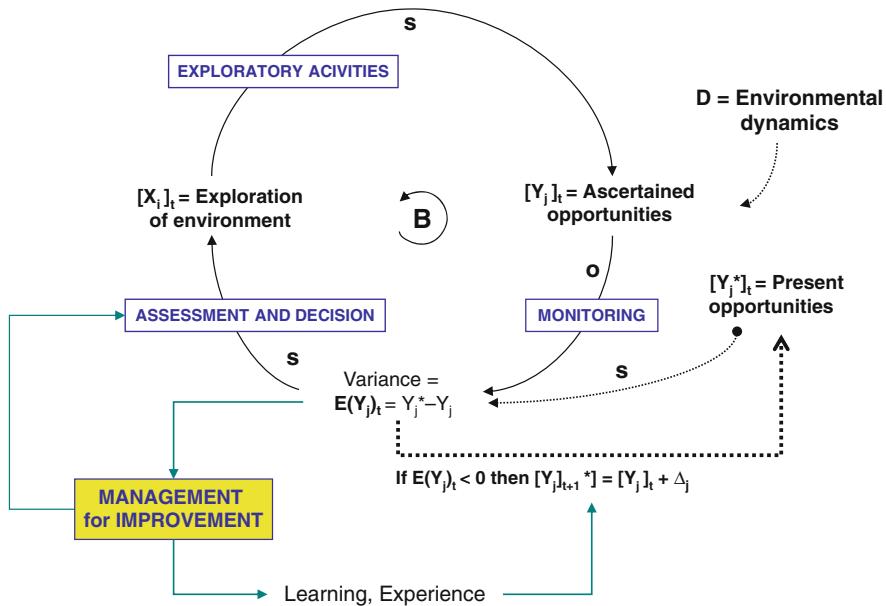
Fig. 8.26 Model of a firm as an exploratory system of improvement

the information becomes. For example, each head of sales for an area writes a weekly report summarizing the data from the salesmen in the area; these weekly reports are passed on to the head of area sales, who summarizes it and transmits the product sales figures to the marketing head, who synthesizes the figures for sales for all products and transmits the results to the managing director, who, together with data for costs, drafts the reports of the gross results for the board of directors.

It is intuitively clear that in order to evaluate and control operational performance at the various organizational levels the data from activities, operations, and processes that produce the economic and financial variables must be known in *real time*, so that the variances with respect to the objectives can be readily perceived and the necessary corrective measures can be taken. In order to carry out control in *real time* the synthetic values for an entire period of observation (for example, production, sales, monthly or quarterly stock) are not particularly significant, since they do not permit us to observe the dynamics regarding the underlying phenomena to be controlled (sales, receipts, and progressive losses), thus making it difficult to readily perceive the abovementioned variances.

For this reason the technique of constructing *dashboards* (or *management cockpits* or *tableau de bord*) to detect in real time the dynamics of the fundamental variables to control continues to spread (Mella 2012). For example, in hypermarkets dashboards are used to highlight the flows of the most relevant types of goods: stocks, resupplies, average supply times, average times goods are exposed on shelves, losses/damage, thefts, etc. In large hospitals dashboards are created for
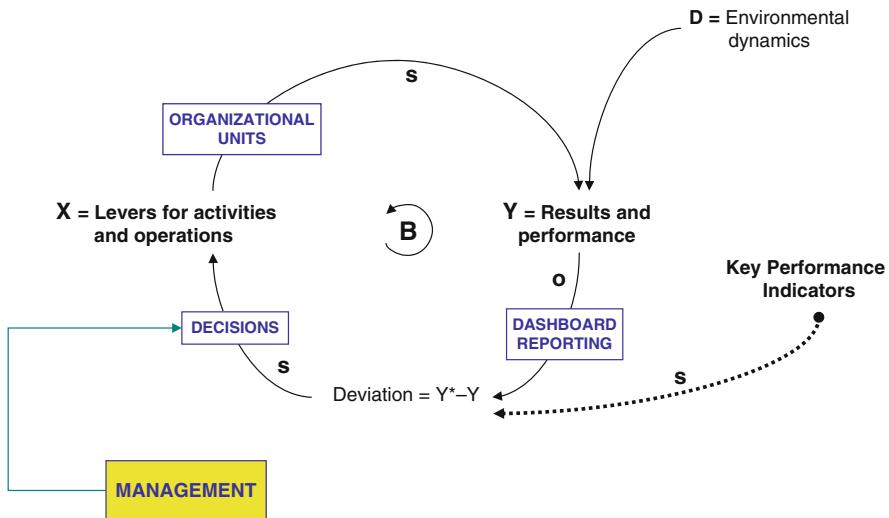
Fig. 8.27 The role of dashboards in performance control

the daily control of the operational dynamics of the main clinics or divisions: numbers and types of patients admitted, length of waiting lists, average hospital stay, type of treatments, percentage of failures, multiple interventions, average amount of drugs administered, average consumption of medication materials, etc. In control theory, *corporate dashboards* must be considered in all respects as *continual reporting* instruments for monitoring a system of *performance objectives* and *standards* in order to allow for the control of operations and personnel at a specific operational level. The dashboard data can also be generally defined as *key performance indicators* (KPI), since they monitor the *performance* regarding the objective of the control process, as illustrated in the *Ring* in Fig. 8.27.

There is no set rule for constructing effective dashboards, even though several principles should be considered, such as:

– Creating above all an organizational culture of action aimed at achieving performance
– Integrating this performance culture with one characterized by an acceptance of performance control
– Creating dashboards to monitor the specific performance of certain personnel and defined functions (*key performance areas*)
– Including in the dashboard only those (few) KPIs that represent realistic performance objectives for the area in question
– Graphically representing the KPIs in an intelligible manner, using different characters, colors, and graphs to highlight them so that they can be immediately utilized

Fig. 8.28 An example of a dashboard (*source*: http://www.projectmanager.com/project-dashboard.php

– Making clear, even graphically, the systematic links among the various KPIs, avoiding the use of colors and formats that can mislead perceptions of the important information for which the dashboard was designed in the first place

Figure 8.28 presents an example of an attractive graph of an organizational dashboard for top management.

8.9.2   The Balanced Scorecard in Performance Management Control Systems

Management performance control systems cannot be limited to merely controlling financial performance. A number of other interesting variables can serve as performance indicators. The *balanced scorecard* (BSC) is one type of *corporate dashboard*, created by Kaplan and Norton (1992, 1996), which provides top management with information for a continual evaluation of the performance of an entire firm. The BSC considers four *strategic variables* or *perspectives* (or focus) held to be fundamental for providing management with a "balanced" perspective of the *strategic performance* of the firm and an evaluation of the progress toward the objective of *shareholder value* creation imposed by *shareholders*.

1. *Financial perspective*: how the organization wishes to be viewed by its shareholders.

Perspective	Number of Measures	Weight
Financial	5	22%
Customer	5	22%
Internal Business Processes	8 to 10	34%
Learning & Growth	5	22%

Fig. 8.29 Weights and measures of the BSC perspectives (*source*: Kaplan and Norton 2001, p. 375)

2. *Customer perspective*: how the organization wishes to be viewed by its customers.
3. *Internal business processes perspective*: through which processes must the organization develop its abilities in order to satisfy its shareholders and customers.
4. *Organizational learning and growth perspective*: which changes and improvements must the company make to implement its vision.

As a corporate dashboard designed to monitor corporate strategy, the BSC presents three unique features: it includes among the KPIs both financial measures and nonfinancial ones that can be adapted to the type of firm implementing them; it is directed mainly at top management and updated relatively slowly. The "balance" of the scorecard is reflected by the balance between outcome measures and performance driver indicators and that between financial and nonfinancial measures. In short, each perspective is represented in a *scorecard* and assigned a weight of relative importance; for each perspective a limited number of *performance measures* managers hold to be truly significant are included in the BSC, as shown in Fig. 8.29.

As an instrument of strategic control, the role of the BSC can be represented by a model entirely similar to the one in Fig. 8.30.

The following measures are particularly efficient in choosing each perspective:

(a) Measures for the financial perspective: value of the action, growth in profits, profit rate, ROI, EVA, ROE, operating costs, operating margin, corporate objectives, survival, profitability, growth, cost reduction, increase in ROI, cash flow, earnings, increase in earnings, profit rate of shares, and so on (Mella 2005b).
(b) Measures for the client perspective: service level, market share, new clients, new products, new markets, customer satisfaction, customer loyalty, product reliability, perceived quality of the product and/or collateral services, customer complaints, etc.
(c) Measures for the internal perspective: increase in efficiency, quality of processes, utilization rate of production capacity, stock storage period, waste, recycling rate of production waste, remanufacturing, lead time, average unit cost, employee morale, motivation, and so on.
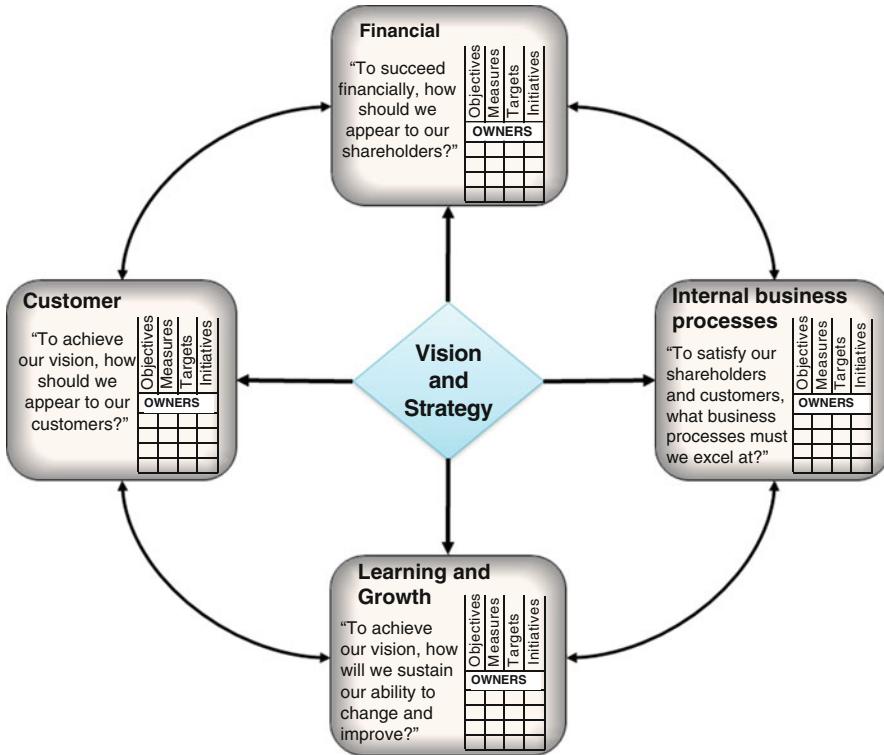
Fig. 8.30 The BSC as a strategy-forming instrument (*source*: Mella 2012: p. 190)

(d) Measures for the learning and growth perspective: trend in value creation, product diversification, supplier diversification, increase in R&D, risk diversification, strengthening of internal control, development of new products, continual improvement, technological leadership, employee involvement, etc.

According to Kaplan and Norton, these perspectives and measures can, in particular contexts, be supplemented by others or enriched by new measures:

> The four perspectives could be viewed as a scheme of reference and not as a straightjacket. Many organizations use the BSC and establish relative weights for each of the scorecard measures. These relative weights are used to evaluate performance (Kaplan and Norton 1996, p. 34).

> If, as we have indicated, the scorecard could guide us in growing our business, then it is natural to believe it possible to change the number of perspectives, areas, or focusses (Olve et al. 1999, p. 120).

If also used as an instrument for determining strategy and not only for monitoring its implementation, the BSC takes on the dual role of an objective and a way of obtaining measurements, thus allowing the organization to determine the variances between the chosen strategy and the actual results.

8.9.3 Control of Projects Through Multi-objective Control Systems

All organizations have projects, more or less complex, but some are specialized in the production of projects by contracting out and includes:

– The construction of specific and unique products: for example, publishing a book, making a film, constructing a machine, and setting up an assembly line
– The production of large works (construction) such as skyscrapers, stadiums, bridges, ships, and missiles
– The study of a prototype and the launch of its production
– Undertaking an advertising or a marketing campaign

Every project requires large investments which are affected by the length of the project and the quality of the work to be done. Therefore, a project must be carefully controlled from three points of view (Mella 2012):

1. *Quality efficiency* in order to guarantee that it is appropriate for use by the client (functional objective).
2. *Time efficiency* in order to *verify* that the project is carried out in the minimum time allowed, given the technical length of the various activities, the user's needs, and the constraints from the producer's equipment capacity.
3. *Economic efficiency* to *verify* that the project is achieved with the minimum cost, given the efficiency objectives and the coordination needs regarding other projects under way at the same time.

When the organization must carry out complex projects, with a precise starting and ending date, it turns to special control instruments, among which *scheduling and time programming*. In order to control and optimize projects, various techniques have been developed to represent the projects, the most widespread of which are the Gantt diagrams, developed by Henry Gantt around 1910—in order to build sequential models, or *chrono programs*, of the project; the CPM, developed in 1957 by Du Pont to solve scheduling problems; and the PERT, developed in 1958 by the US Navy for the Polaris missile program.

The latter two techniques build *network models* of the project, or *activity maps*, which provide a view of the links between the activities rather than their chronological order (Woolf 2007). Today these three techniques are joined in a single *graph technique* called PERT/CPM, which is integrated by the Gantt diagrams; this grid is produced by Microsoft Project™ or other software readily found on the Internet (http://www.criticaltools.com/pertchartexpertsoftware.htm) (Kerzner 2003).

The control of projects through CPM and PERT is based on the idea that the duration of an activity, which requires the use of resources over time, depends on the quantity of resources needed to carry out the activity, and thus on the cost, according to an inverse relationship. In order to reduce the duration it is necessary to increase the quantity of resources per unit of time; vice versa, a reduction in
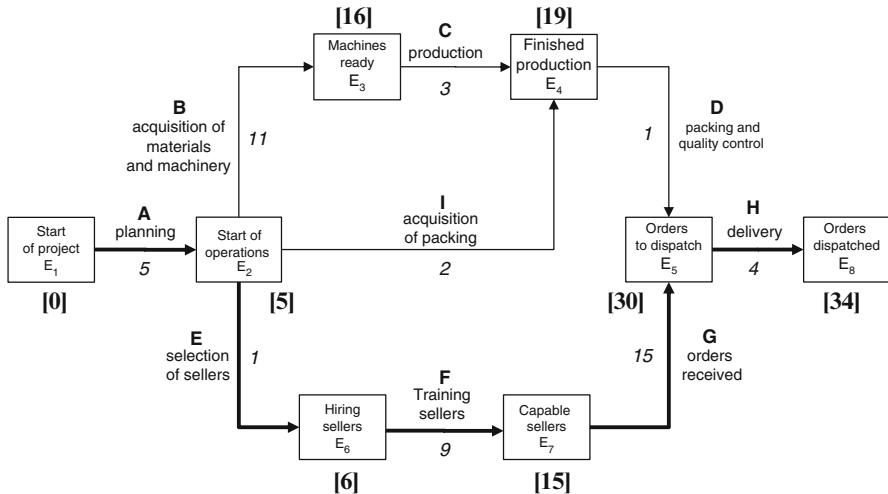
Fig. 8.31 The network for a product-launching project

resources leads to an increase in the duration of any activity. It is thus possible to control the *duration* by modifying the *cost* or to control the *cost* by increasing the activity's *duration*. The basic difference between these two techniques is in the method for estimating the duration of the activities. CPM produces a deterministic estimate, while PERT a probabilistic one, identifies three durations, pessimistic, normal, and optimistic, and takes their weighted average, with the normal estimate having a weight four times that of the pessimistic and optimistic ones.

Very briefly, *temporal* economic control entails the following steps:

1. Analyzing the activity to carry out
2. Building the network of activities for the project, which is a *model* that presents in a reciprocal relationship the various activities that constitute the project, precisely specifying the *order relationships* among the activities themselves and determining the *initial* and *final* event of the entire project
3. Estimating the length of time and cost of the various phases
4. Calculating the length of time and cost of the entire project
5. Optimizing the activities in order to reduce the length of time and costs
6. Scheduling the implementation of the activities (Gantt diagrams)
7. Optimizing the workloads of the personnel and the machines

Figure 8.31, which represents the *map* of the phases (in summary form) of a project to launch a new product, allows us to observe the basic elements of grid planning (this example is in Mella 2012):

(a) The eight blocks—indicated by E_1, E_2, ... E_8—show the *events* that represent the states of progress of the project.

(b) Each event (state of progress) is marked by a *date* that places it sequentially with respect to the events that precede it; by convention, the *initial event* is assigned the date [0].

(c) The nine arrows—indicated by A, B,... I—represent the (macro) activities that produce progress in the implementation (state of progress) of the project.

(d) Each arrow (activity) is characterized by a length of implementation—in specific time units (days, weeks, months, etc.)—written in italics next to the name of the activity.

(e) The date of each event (state of progress) of the project (block) is determined by the sum of the durations of the activities, the succession of which forms a *chain* that, starting from the initial event, leads to that event.

(f) When the same event is produced by several *chains* of activities, then the date of the event is determined by the duration of the longest chain that reaches it; for example, E_4 is dated [19] since the date is determined by the longer of the two chains that reach it: [A → B → C], with duration 19, and [A → I], with duration 7.

(g) Through this procedure we determine the date of the *final event*, which represents the date for the conclusion of the *entire project*, that is, its duration.

(h) The chain of activities (shown by the thicker arrows) that determines the final date—in this case, [A → E → F → G → H], with duration [34]—is called the *critical path*, since lengthening the execution by only a single time unit of one of its activities would lengthen the overall length of the project; vice versa, to reduce the duration of the project, the length of execution of some critical activities is necessary.

(i) The execution of the *critical activities*, which form the critical path, must thus be specially controlled.

(j) It is clear that the activities that lie on the non-critical chains are subject to possible *shifting* as regards the initial and ending dates and, above all, the *length of* execution; PERT defines as *slack* or *float time* the amount of time that a *non-critical activity* can be delayed without causing a delay in the entire project; the *non-critical activities* can be carried out more slowly, thereby reducing their float time and providing them with fewer resources per unit of time.

(k) For such simple projects the calculations can be carried out in an equally simple manner. However, PERT/CPM provides concrete algorithms to automatically quantify length, costs, and slack times, even for projects with hundreds or thousands of activities.

With knowledge of the resources used for the various activities, the consequent length of the activities, and the qualitative standards of the work undertaken, it is possible to control the project through a *system* of control systems that reflect the project model itself. For each activity a *micro* control system is developed to monitor the *length, cost,* and *quality* of execution; a broader system concerns the various *chains* in the project, and with even broader synthesis we can construct the *macro* control system for the entire project.
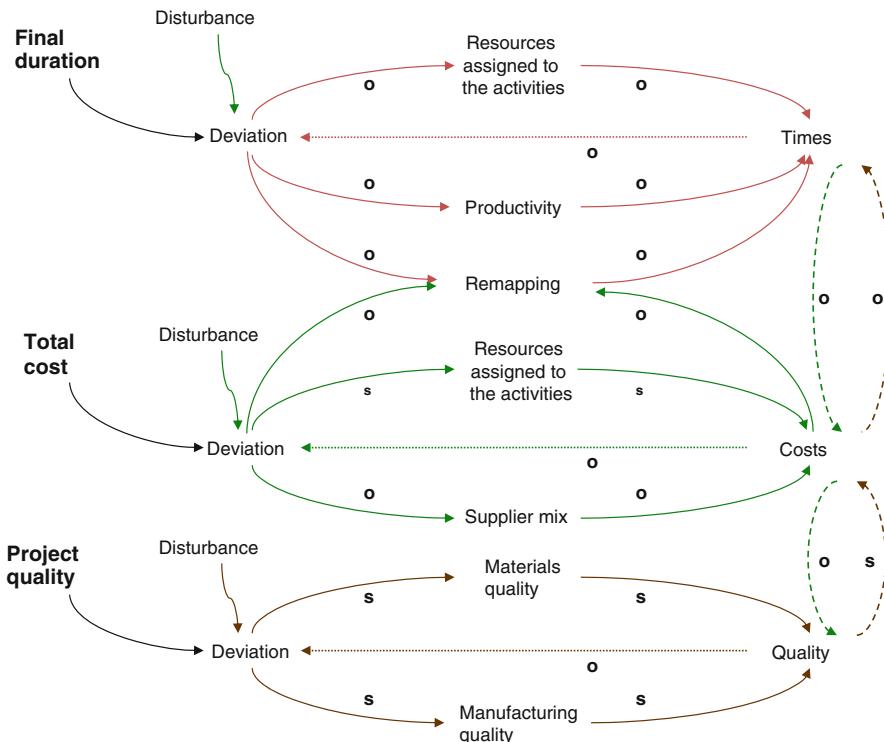
Fig. 8.32 Multi-objective system of control for projects

This *macro* control system is not only multi-lever (Sect. 4.5) but also multi-objective in a true sense (Sect. 4.6), since the contracting firms must verify *contemporaneously* the achievement of the objectives regarding *length, cost,* and *quality* agreed to with the commissioning firm; usually, the control of one of these objectives interferes with that of the others.

As can be seen in the project in Fig. 8.32, a shortening of the overall duration can lead to greater cost and often to a decline in the level of quality.

Thus, it is fundamental to prioritize the objectives. In typical business projects all three objectives are relevant, and a *policy* establishing the priorities must be set at the beginning. To carry out the control, several classes of control *levers* can be used, which involve:

1. Varying the quantity/quality of the resources assigned to each activity in order to control the duration and quality of the critical and non-critical activities (operational lever)
2. Varying the starting date of the activities, especially the *non-critical* ones, by optimizing the *machine load* and *capacity* needed for manufacturing, with appropriate *algorithm assignments* (operational lever)

3. Acting on the *float times* and on the *non-critical activities* in order to free up resources for the *critical* activities (operational lever)
4. Selecting appropriate suppliers and productivity increases in order to control the costs and quality of execution (extraordinary lever)
5. Redesigning the structure of the project (remapping) in order to modify the *critical* and *non-critical* chains, thereby eliminating or merging certain activities (structural lever)

If the project extends over a vast horizon (activities carried out in various places) and a lengthy one (long duration), then there is greater probability that external disturbance events can alter the time, cost, and quality of the activities. When such events cause errors that cannot be eliminated through the control levers, it is necessary to recognize the impossibility of achieving the original objectives and to prepare a *revision* of the project for those parts that must still be achieved.

8.9.4   Model of a Control System of a Warehouse Based on Wilson's Formula

Impulse control systems, which we examined in Sect. 4.3, can be used to control the supply of goods, materials, and components in production organizations based on the general model of control systems for warehouses illustrated in Fig. 8.22. This is not a technical control of stocking and un-stocking, which depends solely on the size of the warehouse, since it is necessary to also introduce the costs relative to the storage system. These costs come from two sources:

(a) The *supply cost* for a single shipment, which is assumed fixed, whatever the size (for example, 100 USD for each order).
(b) The average *storage cost* per unit of stock (for example, 2 USD per day per piece); the overall storage costs will thus be proportional to the total value of the average stock, $q/2$, which arises from the shipment purchase.

The unique problem of such a control system is to calculate the optimal supply shipment (economic order quantity: EOQ), that is, the quantity to order that will ensure the minimum storage cost, which is the sum of the supply costs and stock maintenance costs (Waters 1992; Woolsey and Maurer 2005). There are several methods to calculate EOQ, but the easiest is to use Wilson's formula in its simplest version (Wilson 1934), where it is assumed that the value of stored goods has no influence on costs; in other words, the purchase price is not included in the calculation; for example, Arsham (2006) and Woolsey and Maurer (2005) present a vast collection of calculation modules based on assumptions that are alternatives to Wilson's formula.

We use the following symbols:

- d, the daily need of stored goods
- T, the reference period (for example, 250 work days)
- $Q = dT$, the total (or annual) need for period T
- EOQ, the unknown amount of supply shipment
- $n = Q/\text{EOQ}$, the number of shipments requested during the year
- ca, the supply cost of a single shipment
- cs, the average cost of storage per unit of stock
- $t = \text{EOQ}/d$, the length of stock at the end of which a new order must be made

Some elementary mathematical operations give us Wilson's formula for the calculation of EOQ based on the simple assumptions we have made:

$$\text{EOQ} = \sqrt{\frac{2\text{ca}Q}{\text{cs}}} = \sqrt{\frac{2\text{ca}}{\text{cs}}} * \sqrt{Q}$$

This expression immediately shows that the amount of EOQ increases less than proportionally to the increase in the annual need for Q. Once EOQ is known we can determine the other elements useful for the control, in particular the number of shipments to order and the average duration of a shipment, during which we get the daily output, "d," and at the end of which we need a new input. Figure 8.33 provides a numerical demonstration of the application of the warehouse control based on the data inserted in the initial control panel. Applying Wilson's formula, the optimal shipment is EOQ = 1,000 units. Since the daily consumption is $d = 40$ units, the warehouse must be resupplied every $t = 25$ days.

We can also introduce the assumption that the orders made on a given day will arrive with a constant delay, "r," and that, in order to avoid delay risks, the warehouse manager desires a safety stock, ss $= rd$. If, for example, the delivery delay was on average equal to $r = 5$ days and $d = 40$ pieces per day, then a safety stock of $Y^* = 200$ pieces would be needed. For the warehouse manager this simply means moving up the initial supply by 5 days with respect to the shipment duration period, $t = 25$ days. Since we have assumed a safety level of $Y^* = 200$ units, the initial supply takes place on day 20 instead of day 25. Figure 8.33 also presents the stock dynamics (continuous consumption and discontinuous supply) for the entire duration of the control period, which is assumed to be equal to $T = 250$ work days.

d = daily need		40	ca = supply c.	100.00	d = daily need	40.00
T = days in the year		250	cs = storage c.	2.00	riordino Y_0 =	1000.00
Q = total need for period		10000	r = delay	5.00	safety stock Y^* =	200.00
					t = length of stock	25.00

$$EOQ = \sqrt{\frac{2\,ca\,Q}{cs}} = \sqrt{\frac{2\,ca}{cs}} * \sqrt{Q}$$

t	Δ X1	Δ X2	Δ Y	Y* = objective	Δ objective	E = disturbances	E(Y) = Y* - Y
0	0.0	0.0	1000.0	200.0	0.0	0.0	-800.0
1	-40.0	0.0	960.0	200.0	0.0	0.0	-760.0
2	-40.0	0.0	920.0	200.0	0.0	0.0	-720.0
3	-40.0	0.0	880.0	200.0	0.0	0.0	-680.0
4	-40.0	0.0	840.0	200.0	0.0	0.0	-640.0
5	-40.0	0.0	800.0	200.0	0.0	0.0	-600.0
6	-40.0	0.0	760.0	200.0	0.0	0.0	-560.0
7	-40.0	0.0	720.0	200.0	0.0	0.0	-520.0
8	-40.0	0.0	680.0	200.0	0.0	0.0	-480.0
9	-40.0	0.0	640.0	200.0	0.0	0.0	-440.0
10	-40.0	0.0	600.0	200.0	0.0	0.0	-400.0
11	-40.0	0.0	560.0	200.0	0.0	0.0	-360.0
12	-40.0	0.0	520.0	200.0	0.0	0.0	-320.0
13	-40.0	0.0	480.0	200.0	0.0	0.0	-280.0
14	-40.0	0.0	440.0	200.0	0.0	0.0	-240.0
15	-40.0	0.0	400.0	200.0	0.0	0.0	-200.0
16	-40.0	0.0	360.0	200.0	0.0	0.0	-160.0
17	-40.0	0.0	320.0	200.0	0.0	0.0	-120.0
18	-40.0	0.0	280.0	200.0	0.0	0.0	-80.0
19	-40.0	0.0	240.0	200.0	0.0	0.0	-40.0
20	-40.0	0.0	200.0	200.0	0.0	0.0	0.0
20	0.0	1000.0	1200.0	200.0	0.0	0.0	-1000.0
21	-40.0	0.0	1160.0	200.0	0.0	0.0	-960.0
22	-40.0	0.0	1120.0	200.0	0.0	0.0	-920.0
23	-40.0	0.0	1080.0	200.0	0.0	0.0	-880.0
24	-40.0	0.0	1040.0	200.0	0.0	0.0	-840.0
25	-40.0	0.0	1000.0	200.0	0.0	0.0	-800.0
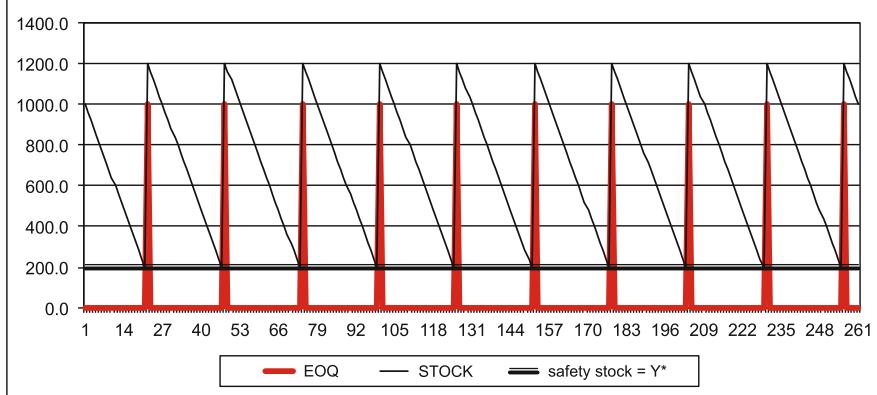
Fig. 8.33 Impulse control system of a warehouse using Wilson's formula for calculating the optimal reordering shipment

8.10   Summary

After having identified the innumerable *Rings* that characterize individual behavior in our environment (Chap. 7), and *the* Rings that regulate the behavior of biological populations in ecosystems, collectivities, and societies (Chap. 8), I have tried to "see beyond" by examining control systems that influence *organizations*.

In particular, I have examined the *Rings* that allow organizations to exist:

(a) As autopoietic systems, along the lines envisioned by Maturana and Varela (Sect. 8.1, Fig. 8.2), and as teleonomic systems, following Monod's approach (Fig. 8.7)
(b) As vital systems, following Beer's VSM (Sect. 8.2, Fig. 8.8)
(c) As systems for efficient transformation, following the MOEST approach (Sect. 8.3, Fig. 8.10)
(d) As systems requiring management control (Sects. 8.4 and 8.5, Fig. 8.14)—at both the strategic (macro control) and operational (micro control, Sect. 8.7) through the implementation of management control systems and performance control systems
(e) As cognitive, intelligent, and exploratory systems (Sect. 8.8)

We have examined several basic instruments for micro management control (Sect. 8.7):

1. Stock control (Fig. 8.22) and a control system of a warehouse based on Wilson's formula (Fig. 8.33)
2. Labor productivity control
3. Control of the various forms of quality (Fig. 8.23)
4. The use of dashboards in performance control systems (Sect. 8.9.1, Fig. 8.27)
5. The use of the balanced scorecard in performance control systems (Sect. 8.9.2, Fig. 8.30)
6. Project control through the PERT/CPM/Gantt technique (Sect. 8.9.3, Fig. 8.32)

Chapter 9
The Magic Ring Explores Cognition and Learning

I have said that what gets from territory to map is transforms of differences and that these (somehow selected) differences are elementary ideas. But there are differences between differences. Every effective difference denotes a demarcation, a line of classification, and all classifications are hierarchic. In other words, differences are themselves to be differentiated and classified. In this context I will only touch lightly on the matter of classes of difference, because to carry the matter further would land us in the problems of Principia Mathematica (Gregory Bateson 1972, pp. 463–464).

I shall reconsider human knowledge by starting from the fact that we can know more than we can tell. This fact seems obvious enough; but it is not easy to say exactly what it means. Take an example. We know a person's face, and can recognize it among a thousand, indeed among a million. Yet we usually cannot tell how we recognize a face we know. So most of this knowledge cannot be put into words (Michael Polanyi 1967, p. 4).

This chapter concludes the journey we have made until now by examining Control Systems that make learning and knowledge possible. Since the topic of knowledge is incredibly vast and cannot be scaled down to fit in with the content and size of this book, I have limited the examination to those *Rings* needed for the conscious "mind" to function. I have deliberately chosen to use as a guide Gregory Bateson's model of the "mind" as a calculator of various levels of differences, undertaking to formally operationalize this model. I hope to succeed in presenting the reader with a convincing model of how knowledge is formed through the perception and systematization of differences. Knowledge which is structured "in" the mind can only be revealed through some behavior produced by the cognitive system. I have therefore examined the problem of the formation of signs and the use of language to reveal and communicate knowledge. I have only touched on the problem of the relation between knowledge and the "brain" and the problem of conscious knowledge. The topic of learning is only examined with regard to several aspects that involve the construction of models and the explanation process. Once again, the action of *Rings* is fundamental to the development and transmission of knowledge. I invite the reader to read the initial sentences in Sect. 9.1, in which I point out the difficulties in dealing with the topic of knowledge by using mainly the sense of sight and employing a written language.

9.1   Dimensions of Knowledge

This chapter will demonstrate that Control Systems are indispensable for the "knowledge process" and for the formation of the "content of knowledge" through which "cognitive individuals" are structurally coupled to the external environment, trying to survive by reacting to disturbances.

It is not simple to analyze in a book, "as humans," the knowledge phenomena in living beings equipped with a nervous system. The *first difficulty* is entirely clear: dealing with knowledge through a book means speaking with a *written language* (which belongs to a specific linguistic code (English)) which represents the result of a highly complex and evolved mental activity in human beings. Using this language may make it difficult to describe the elementary mental processes that precede the production of the language; for this reason I shall try to integrate the lexical language (written English) with more abstract graphic models. The *second difficulty concerns* the almost obligatory use of examples that make use of the *sense of sight*, thus making it difficult to speak and write about stimuli and sensations which derive from other senses without being able to represent them nonlinguistically. While I can say: "Observe these figures. . ..," it is difficult for me to say: "Try these flavors . . . or smell these scents. . .." Flavors and scents must be indicated through words; this assumes the reader already has knowledge of their meaning.

The *third most important difficulty* is *methodological in nature*, since, in my view, there are at least four dimensions for observing mental phenomena, all of which are interconnected, even though they can be examined separately:

1. The "brain" dimension (in general, the central nervous system) as a "physical system" that produces basic neuronal processes from which knowledge is built as the computation of neuronal stimuli. The connections among the neurons and the various types of schema that occur through the ordering of combinations of perceptions produce the mental sensations corresponding to colors, flavors, smells, etc. Such stimuli are not in the environment and the brain does not perceive them as the result of specific stimuli; they derive instead from the interconnection of stimuli from multiple sensors operating in parallel, which are combined to produce the variety we perceive. Simulation through cellular automata seeks to demonstrate the hypothesis of the *connectionist* functioning of the brain:

   The most important idea is that of a node. Nodes are simple energy-transmitting devices, which in the simplest case are characterized at a given time by their degree of activation, or propensity to affect other nodes. Nodes are connected to one another by stable energy conduits, by means of which active nodes tend to alter the activation of other nodes. [. . .] Where are the representations in connectionist models? On one conception, individual nodes (or states of individual nodes) are representations. Perhaps the most famous example is the "interactive activation model" of reading (Rumelhart and McClelland 1982).

   The network contains "word nodes" each of which standardly becomes activated as a result of the presentation of a particular word, and each of which represents the presence of that word. These word nodes are activated by "letter nodes" each of which represents and standardly responds to the presence of a particular letter at a particular position in the word.

Finally, each letter node is activated by "feature nodes" each of which represents and standardly responds to the presence of a certain feature of the shape presented at a particular position: a horizontal bar, a curved top, etc. (Lormand 1991, online).

2. The "mind" dimension, as an *emerging* system and feature of the brain's neuronal activity, which produces knowledge as a process necessary for the survival of living beings; this dimension cannot be directly observed but it can be deduced from observing behavior. Therefore, without behavior we cannot infer the existence of knowledge. As humans we can all directly verify this dimension through our awareness; we know nothing about how our brain functions, with its 100 billion neurons, but we feel heat and cold, we see color, we distinguish signs and try to understand their meaning.
3. The "awareness" dimension, as *conscious* knowledge (thought-containing content, including every form of language), which is a capacity presumably limited to man only.
4. The "intelligence and learning" dimension, as processes in the formation and modification of the content of knowledge.

In a chapter dedicated to Control Systems it is not possible to treat these four aspects of knowledge, which have vast implications in many disciplines. Even a summary examination would be beyond the scope of this chapter. I will thus only present and attempt to operationalize Gregory Bateson's theory on the functioning of the mind (point 2 above), which produces knowledge through the elaboration of differences at multiple levels. My objective is to demonstrate that knowledge, even at the level of the "mind," assumes an incredible number of Control Systems.

The "brain dimension" (point 1) and its relation to the cognitive and behavioral process has been studied by neuroscience.

> Scientists still have not uncovered the full extent of what the brain can do. This single organ controls every aspect of the body, ranging from heart rate and appetite to emotion and memory. [. . .] Neuroscientists specialize in the study of the brain and the nervous system. They are inspired to try to decipher the brain's command of all its diverse functions. Over the years, the neuroscience field has made enormous progress. Scientists continue to strive for a deeper understanding of how the brain's 100 billion nerve cells are born, grow, and connect. They study how these cells organize themselves into effective, functional circuits that usually remain in working order for life (Society for neuroscience, web page).

We do not "observe" knowledge in the brain (more generally speaking, in the central nervous system), but only neuronal phenomena as the condition for knowledge to emerge, in the sense that, according to the rules of implication: "knowledge emerges from a functioning brain" since "if there is knowledge then there is a functioning brain," and "if the brain does not function, then knowledge will not emerge." The brain is conceived of as a "computational machine" that functions as a closed system whose inputs are its own states:

> Some of us consider the Central Nervous System to be a fundamentally closed system with its basic organization geared toward the generation of intrinsic images (thoughts or predictions), where inputs *specify* internal states rather than inform "homuncular vernacles" (Llinás and Paré 1996, p. 4).

This perspective is typical of *functionalism* and *connectionism*, according to which the mind produces processes of knowledge that depend on (are produced by or correlated with) neuronal physical processes. For the neurobiologist Gerald Edelman, who has developed a theory of neuronal group selection, or Neural Darwinism (see the quote below), as a model for brain functioning, ... [The] *mind is a special kind of process depending on special forms of matter* (Edelman 1992, p. 6).But then, Edelman goes on, we can say that mind and matter are both processes, i.e., interactive and not static. Are these processes the same? Or, are they similar to one another; and if they are similar, in what ways are they similar? Or, are they different from one another, so different that conflation would be in grave error? An examination of the vast and well-developed field of neuroscience, which is, moreover, highly technical and specialized, goes beyond the limits of this book; however, I would note that Thomas Henry Huxley had already, in the second half of the nineteenth century, and even without any sophisticated neuroscientific knowledge, observed the relations between the nervous system and behavior with his famous "example of the frog," which led him to conclude (erroneously from a contemporary perspective) that animals are, or behave as if they were, automata:

> As I have endeavoured to show, we are justified in supposing that something analogous to what happens in ourselves takes place in the brutes, and that the affections of their sensory nerves give rise to molecular changes in the brain, which again give rise to, or evolve, the corresponding states of consciousness. Nor can there be any reasonable doubt that the emotions of brutes, and such ideas as they possess, are similarly dependent upon molecular brain changes. Each sensory impression leaves behind a record in the structure of the brain—an "ideagenous" molecule, so to speak, which is competent, under certain conditions, to reproduce, in a fainter condition, the state of consciousness which corresponds with that sensory impression; and it is these "ideagenous molecules" which are the physical basis of memory. It may be assumed, then, that molecular changes in the brain are the causes of all the states of consciousness of brutes (Huxley 1874, pp. 239–240).

If we assume the *connectionist hypothesis* is valid, then the development of knowledge by the functioning brain through the network of its connections is simulated by computational techniques that rely on multiple models of neural networks:

> The computational approach to these levels, which we call functional reconstruction of neuronal systems, strives to perform realistic simulations while giving adequate representation to the particularities of neuronal morphology, the various voltage-dependent and chemically mediated ion currents in membrane-bound channels, neurosecretory activity [...] Flexibility and versatility are the keynotes to this approach which consequently uses a compartmental model of the neuron (Perkel 1993, p. 40).

> The large gaps in our understanding of biological memory schemes and mechanisms, the limited applicability of cellular mechanisms and of connection-modification rules in hardware devices, together with the emergence of feature-detection properties in parallel distributed networks constitute the greatest challenge for computational neuroscience (ibidem, p. 44).

The "awareness" dimension (point 3 above) is encompassed in the field of study of various disciplines (Neuman 2003), first and foremost psychology, which

investigates conscious and unconscious mental processes, the latter, in my view, deriving from nonconscious, innate, or acquired knowledge. According to Gerald Edelman:

> Primary consciousness may thus be briefly described as the result of the ongoing discrimination of present perceptual categorizations by a value-dominated self-nonself memory (Edelman 1989, p. 102).

> My main focus was on perceptual categorization as it related to memory and learning. I proposed that these functions could be understood in terms of "neural Darwinism"—the idea that higher brain functions are mediated by developmental and somatic selection upon anatomical and functional variance occurring in each individual animal (ibidem, p. xvii).

> I proposed that this ability depended critically on two of the most striking features of the brain, its variability and its re-entrant connectivity (ibidem, p. xviii).

The "intelligence and learning" dimension (point 4 above) is vast and includes the formation of experience, voluntary or guided, even in living beings without a nervous system but still coupled to the environment, and the formation of thought, linguistic and nonlinguistic content, which is needed to improve the existence of cognitive individuals. Various disciplines operate in this field of research, among which cognitivism, pedagogy, psychology, and artificial intelligence. "Artificial Intelligence is one of the newest fields in science and engineering. Work started in earnest soon after World War II, and the name itself was coined in 1956" (Russell and Norvig 2009, p. 4)

> Artificial intelligence is a rapidly evolving field of engineering with an ultimate objective to build machines capable of acting and thinking like human beings. The early phase of AI was concerned with developing programs for theorem proving and game playing. Modern AI encompasses various tools and techniques for humanlike reasoning, learning, planning, language and pattern recognition. Artificial intelligence is probably one of the most successful branches of a broad area of computing (Kumar 2008, p. 1).

The next few *sections* will thus be limited to a treatment of the "mind" dimension and the "intelligence" dimension, not so much to describe the content of mental and cognitive processes but to demonstrate how basic mental processes are based on a multitude of *Rings* acting so quickly as to remain entirely hidden from our awareness.

9.2   Bateson's Model: The "Mind" as a Calculator and Transformer of Differences

In his excellent book, *Mind and nature: A necessary unity*, Gregory Bateson proposes an epistemological theory of knowledge based on the simple yet convincing model of "mind" as the capacity of a cognitive system, or individual, to form a representation (map) of the world (territory) through the perception and ordering, even at successive levels, of differences.

1. Mind is an aggregate of interacting parts or components. 2. The interaction between parts of mind is triggered by difference and difference is a non-substantial phenomenon not located in space or time. 3. Mental process requires collateral energy. 4. Mental process requires circular (or more complex) chains of determination. 5. In mental process the effects of difference are to be regarded as transforms (that is, coded versions) of the difference which preceded them. 6. The description and classification of these processes of transformation discloses a hierarchy of logical types immanent in the phenomena (Bateson 2002, p. 92).

Bateson does not take into consideration the brain, but views the *mind* as the "processor" of knowledge, an open system that acts as a "machine"—with multiple inputs—capable of calculating differences, memorizing and comparing these, and finding analogies. Moreover, Bateson—by adopting a simple metaphor—distinguishes between knowledge and what is known, comparing knowledge to a map, what is known to a territory: "The map is not the territory, and the name is not the thing named" (Bateson 1979, p. 30; see also Mella 2012, Sect. 2.12). The map—that is, knowledge—is formed by taking account of the differences the observer perceives in the territory represented; these differences and their transforms are "elementary ideas . . . and these differences are themselves to be differentiated" (Bateson 2000, p. 463). Our "mind," as a "processor" of knowledge, carries out a continual function of discovering relationships in the patterns of differences, and this process leads to the emergence of a hierarchy of differences based on which all knowledge is constructed (Bateson 2000, pp. 454–471; Bateson 2002, p. 106).

Even if Bateson does not explicitly state this, I suggest we recognize that the "mind" is not a logical-abstract entity but the principal source of knowledge for *cognitive systems*, in particular "*conscious cognitive systems*." I define the latter as autopoietic systems which, through the mind, succeed in giving *significance* to *environmental stimuli*, transforming these into reactions (which an observer describes as macro behavior or processes interacting with the environment) in the *search* for positive stimuli for survival, or, alternatively, the *flight* from negative stimuli. In order to produce Bateson's model we must assume that cognitive systems must possess:

(a) *Sensor organs* that can perceive differences (in the form of stimuli, impulses, disturbances, etc.).
(b) *Computational* capacities, through which differences are ordered and these differences themselves are differentiated.
(c) *Memory capacity,* which reconstructs or preserves the computed differences through changes in state (of greater or lesser duration) of the internal processes of memory.
(d) The *evaluation capacity,* needed to give "weight" to the dynamics of the internal states which have favored or hindered the search for positive stimuli or the flight from negative ones.
(e) *Effector organs*, which give rise to environmental behavior involving the search for positive stimuli or the flight from negative ones; more specifically, the effector organs (mimicry, sound emission, signs, etc.) must be able to express thought content.

The mind functions through the action of all the components of the cognitive system, even if Bateson believes *differences* are not substantial phenomena since they do not result from a comparison involving the values of variables the cognitive subject is capable of (must be capable of) determining and comparing in order to produce new differences for subsequent comparison. Rules 5 and 6 above, which define the logic of the mental process, clearly show that Bateson views the "mind" as operating in a holonic fashion by forming a holarchy, that is, "a hierarchy of logical types immanent in the phenomena" (Bateson 1979, p. 122).

To apply Bateson's model we must postulate that, in order to undertake a recursive process to determine successive levels of differences, the "mind-processor of knowledge" must, as a minimum, be "structured" to carry out four basic operations, all of which are necessary:

- *Distinction*, or perception; that is, identifying a *primary difference* with reference to a "null" internal state.
- *Comparison*, which leads to the identification of *differences of differences* by forming *differences* of various levels (through some sense organ), and their inclusion in various types of *dimensions*.
- *Perception of analogies* (through some form of memory), which leads to the identification of ranges of variation between *differences* held not to be "too different" from each other.
- *Association,* which leads to considering a certain number of *differences* as a "different unit," as if they constituted a single *object* of observation.

> In fact, wherever information—or comparison—is of the essence of our explanation, there, for me, is mental process. Information can be defined as a difference that makes a difference (Bateson 2002, p. 91).

As *rule* 5 of "Bateson's conception of mind" clearly indicates, the differences are produced at various levels. Figure 9.1 shows the *Ring* needed to produce the *primary differences* in a cognitive system with a *sensor* capable of distinguishing between the outside disturbances. Referring to Fig. 9.1, I shall define as *basic,* or *primary differences* those differences formed by the cognitive system when the internal mental state is "null" (this is indicated by $s_n =$ "Ø") and the environment produces a disruption or disturbance ($e_n =$ "y") to some sensor organ.

The "mind" carries out a basic comparison to determine the primary difference, which I shall indicate as $d_n^0 =$ " y "; the subscript "n" indicates that the *Ring* operates during a general phase "n" of the process, or at the general instant "n" when it perceives the difference $d_n^0 =$ " y. " In other words, the mind makes a comparison with the "null state"; a basic difference is perceived as the error produced by this comparison. The mind should be able to memorize all the differences, even those which are repeated, and distinguish these from the null state; for example, $d_1^0 =$ " j, " $d_2^0 =$ " k, " $d_3^0 =$ " w, " $d_4^0 =$ " k, " $d_5^0 =$ " j, " $d_6^0 =$ " w, " etc.

Obviously, if the cognitive system had numerous *sensors* of the same type, each capable of distinguishing similar differences supposed to exist at the same instant
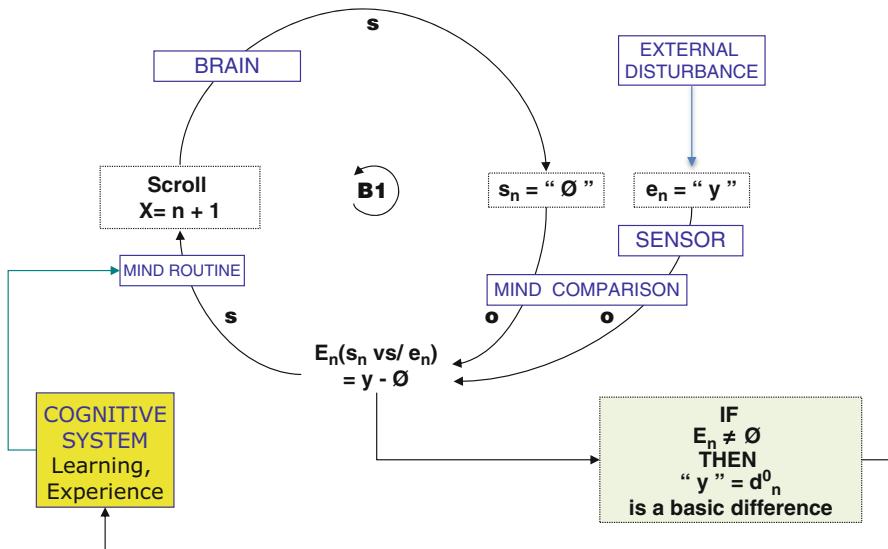
Fig. 9.1   The formation of basic differences through a sensor

(hundreds of contemporaneous perceptions by means of hundreds of spinal hairs on an insect), or had different types of sensors that could distinguish specific differences (visual, auditory, in terms of flavor, etc.), we would have to assume an equal number of *Rings*, similar to those in Fig. 9.1, acting contemporaneously. Adopting "human" terminology, we could say that the cognitive system "sees" light, "tastes" a flavor, "hears" a sound, etc., when it detects a difference with respect to a "null" state.

Figure 9.2 illustrates the *first-level differences* from a comparison of *basic differences* held to be comparable by the cognitive system; for example, various intensities of light, differences in flavor, color, sound, etc. For Bateson, these first-level differences at instant "*n*," which I indicate as d_n^1, can be thought of as a transformation of "zero-level" differences, when the error arising from the comparison exceeds a tolerance threshold, ΔE_n, as shown in Fig. 9.2.

I hold that it is not possible to compare basic differences that derive from different types of sensors: sounds with colors, temperatures with flavors, etc. Colors, flavors, smells, tactile sensations, and so on (expressed in our human language) are examples of basic differences; the color stimuli cannot be compared to flavor stimuli, even if the mind recognizes two stimuli as two colors and two stimuli as two flavors. If the cognitive system has different sensors and *memory*, it can activate a large number of *Rings* to provide the "mind" with a "catalogue" of differences of the same level and an "experience" of perceiving differences. By increasing the level of comparison regarding comparable differences, the mind of the cognitive system can create a "vast catalogue" of differences of various levels which, according to Bateson, can be considered as the result of the transformations
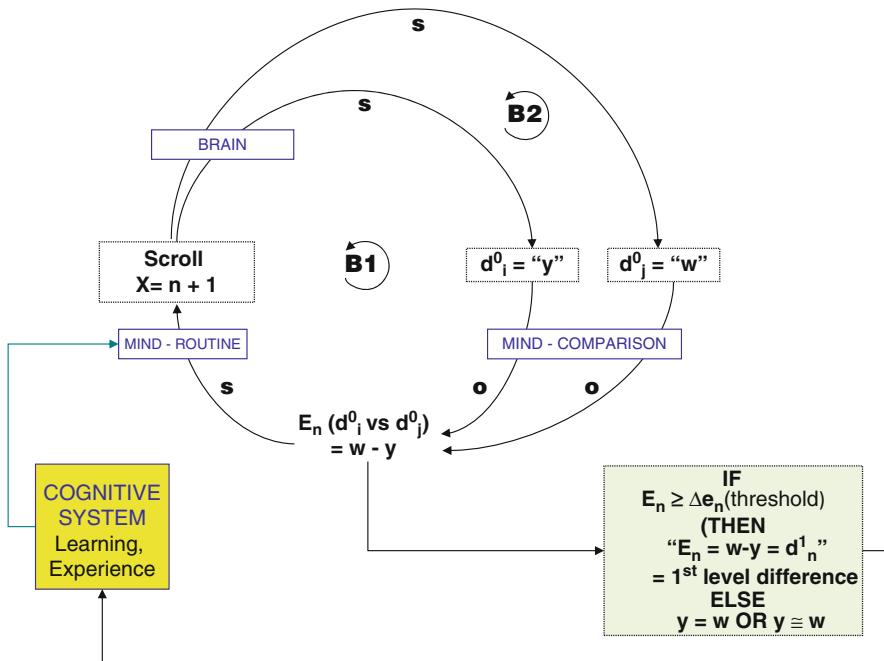
Fig. 9.2 The formation of first-level differences with a sensor

of differences of previous levels through the detection of error by the *Rings* undertaking the comparisons.

Figure 9.2 also illustrates the notion of "equality," which cannot be directly perceived but derives instead from the absence of differences ($E_n = 0$) between differences of the same level, so that there is no significance in identifying differences of a subsequent level. We can all verify this by comparing, for example, two flavors, colors, or sounds; if we do not find any differences we conclude that those flavors, colors, and sound are "equal." If there are minimal differences (below the tolerance threshold), those flavors and sounds are judged to be "entirely similar."

In order to understand how the mind can detect differences at higher levels, it is useful to turn to the idea of *dimensional distance*, which for the mind represents a "*step*" between similar perceptions. If several differences of order zero have been perceived, then a comparison of these can allow different "steps" to be identified which form a dimensional scale of distance. In this way various dimensional categories are formed. The identification of distances can be repeated by in turn considering the distances as perceptions. Thus, higher-order differences are formed, as illustrated, for example, in the following Table 9.1.

Table 9.1 Higher-order differences

Perceptions	Dimensions	Distance	Variation	ETC
Differences of order zero: d_n^0	Differences of order one: D	Differences of order two: Δ	Differences of order three: $\Delta\Delta$	$\Delta\Delta\Delta_k$
Yellow	Color	Chromatic scale	Variations in tone	
Red				
Green, etc.				
Sour	Flavor	Scale of flavors	Intensity	
Sweet				
Bitter, etc.				
High	Height	Scale of heights	Variations in height	
Higher				
Higher still				
etc.				

9.3 Bateson's Model Expanded: The "Mind" as a Calculator of Similarity and Analogy: The Construction of Dimensions

Even if Bateson does not touch on this, I observe that thanks to the *mind's* capacity to compare differences the *cognitive system* can, in addition to detecting higher-level differences or ascertain equalities, carry out two further operations: detecting similarity and analogy.

Two primary differences, "h" and "k" (etc.), though different, can be considered *similar* if they derive from the same sensor organ at a different moment and under different conditions and cannot be compared through primary differences, "v" and "w" (etc.), derived from other sensor organs. For example, [d_n = "green" $\neq d_m$ = "yellow"] and [d_i = "sweet" $\neq d_j$ = "savory"] but [d_n and d_m] cannot be compared to [d_i and d_j], since they derive from the action of different sensors; nevertheless, d_n and d_m are differences which are similar to each other, in the same way that d_i and d_j are similar. I shall define as basic dimensions: $D_A, D_B, \ldots D_M \ldots$, etc., the set of all primary differences, $\{d_{A \cdot n}\}, \{d_{B \cdot n}\}, \ldots \{d_{M \cdot n}\}, \ldots$, etc., the "mind considers *similar* because they can be compared. For example, using our human capacity to assign names to the differences our mind perceives (Sect. 9.10), we can write:

$$D_{COLOR} = \{d_{COLOR\ n} = \text{"green,"} d_{COLOR\ m} = \text{"yellow,"} \text{etc.}\},$$
$$D_{FLAVOR} = \{d_{FLAVOR\ i} = \text{"green,"} d_{FLAVOR\ j} = \text{"yellow,"}\}, \text{etc.}$$

By generalizing, we can assume that after having perceived the primary differences the mind manages, thanks to its *memory*, to order these as distinct dimensional elements. If we assume (again using our human minds) that the mind perceives the primary differences d_r = "blue" and d_s = "humid," it can conclude that:

$$d_r = \text{"blue"} \in D_{\text{COLOR}}$$
$$d_s = \text{"humid"} \in D_{\text{TACTILE}}$$

and that d_r and d_s cannot be compared. Secondly, we can assume that when it perceives a difference, for example, $d_n = $ "dark blue," the mind can immediately determine that "dark blue" belongs to the dimension D_{COLOR}. In human terminology, the mind could recognize that the difference "dark blue" belongs to the differences included in (which make up) the "color" dimension. The mind forms the dimensions as *scales of differences* held to be *similar*. A third hypothesis is to assume that once it has identified the difference $d_{\text{COLOR } q} = $ "dark blue," the mind is also able to compare this difference with other differences contained in D_{COLOR} by identifying second-level differences (light blue, navy blue, grayish, etc.) included in second-level dimensions, such as $D_{\text{SCALES OF BLUE}}$, $D_{\text{SCALES OF RED}}$, etc.

If we assume that the mind, following Bateson's reasoning, can and must function as described above, then we can examine a new, important operation it carries out: *analogy*. *Analogy* is not simply the lack of differences but an "evaluation" of the level of differences; "*v*" and "*w*," though different, are analogous if their differences are considered "too small" to reveal a difference of differences. If ΔE_n indicates the tolerance threshold for judging two differences (at a certain level) which are "not different enough" to justify creating a higher-level difference, then the mind can classify "*v*" and "*w*" as analogous (we can write $v \cong w$, relative to ΔE_n). Only through *analogies* can the mind create concepts, and thus develop knowledge. We can recognize the capacity to identify *differences* and construct *analogies* in a mental experiment proposed by Bateson.

> Let me invite you to a psychological experience, if only to demonstrate the frailty of the human computer. First note that differences in texture are different (a) from differences in colour. Now note that differences in size are different (b) from differences in shape. Similarly ratios are different (c) from subtractive differences. Now let me invite you. . . to define the differences between "different (a)," "different (b)," and "different (c)" in the above paragraph (Bateson 2000, pp. 463–464).

The experiment proposed by Bateson refers to differences that can be perceived by different sensors, and it applies to our human experiences; however, in my view, it can also be valid for much more basic cognitive systems. I propose translating this mental experiment into practice by observing the following signs on a sheet of paper (objects) (I have underlined those letters which are colored and which in printed form appear as black and white):

boy **b o y** Bo**Y** B*O*bO <u>Bo**A**</u> <u>BO</u>a B **O** **a** b o A **A a** B *B* <u>a</u>

In what way can we say that these 30 objects, which can be perceived only through the visual sensor, are equal, similar, or different? The obvious response depends on our capacity to compare them and perceive differences, similarities, or

analogies (absence of "significant" differences) between the various letters or groups of letters (objects). There are thus a number of ways to perceive differences, similarities, or analogies between previously perceived objects; this capacity depends on the mental organs (important here is memorization and the capacity to repeat the observations), but there is no doubt that this capacity consists in the perception of *difference that makes a difference* during the operations of comparison (observe the movement of our eyes when you compare the letters).

In order to understand how the Ring in Fig. 9.2 acts, we need to ask ourselves if the objects {o **o** o *O* O o **O** O} are equal. The answer undoubtedly is "No," considering the differences in font, color, size, thickness, etc. Now let us ask ourselves "Why?" "In which dimensions do they differ?" Are these objects similar or analogous? "Why?" Are the objects {**b** **b** B B B B b *B* *b* B} equal, similar, or analogous? "Why?" Are the objects {o **o** o *O* O o O **O** o} and {**b** **b** B B B b *B* *b* B} similar or analogous? Finally, and we could continue in this way with dozens of other comparisons: are the objects {**boy** **b** **o** **y** BoA Boa} equal? In how many dimensions do they differ? Are they analogous? Why?

Independently of the many answers of "different, equal, similar, or analogous" each of us could give, it is impossible not to note how many times we apply the magic *Rings* in Figs. 9.1 and 9.2. Clearly these basic operations to identify and compare differences in order to determine new differences are possible only if we assume that the *structure of the cognitive system* can set off a large number of superfast *Rings* (magic) operating, for example, as indicated in the model in Fig. 9.1 or 9.2. The capacity of the cognitive system to calculate and compare differences obviously depends on the sensors which are available. Consider how reality has been expanded due to the invention of the microscope or telescope, instruments that, as we know, have increased the eye's capacity to perceive differences.

Once Bateson had affirmed his epistemological principle—the "mind" constructs knowledge through differences—he did not produce an operational-logical process to derive concepts, ideas, and meanings from differences (of differences).

Strictly following the direction outlined by Bateson, I propose three steps:

1. Above all, to make Bateson's definition of "mind" operative by introducing simple symbols, I show how it is possible for any "mind"—not necessarily only the human one—to construct *descriptions* of objects and *definitions* of concepts only by making use of "primitive" operations involving *comparison*, the *identification* of differences, and *analogy*; the symbols introduced give a meaning to the observation, identification, and comparison of objects and concepts, which will allow me to formalize several models or moments that define the concept of knowledge.

2. To apply the same conceptual framework to define the process of *denomination* through which the "mind" manages to represent objects (descriptions) and

concepts (definitions) through *signs* (descriptions of signs) and *signifiers* (definitions of signs), thereby forming *semes*.

3. To apply the symbols and the concepts of *description, definition* and *denomination* to operationally deal with the problem of *truth* as correspondence, making use of a reliable *process of determination*.

9.4   The First Step Toward Knowledge: The Process of the *Description* of Objects

The first model-moment of knowledge is the *description* of an object (material or immaterial; individual, group, or system; phenomenon, event, act, operation, process, etc.) or of a given part of reality conceived, in the broadest sense, as a unit of observation by the cognitive agent.

Let me start by saying that in Bateson's model the mind does not *observe* objects but compose them through operations involving the perception of differences and the search for similarities. A certain number of *differences* at various *levels*, even recognized as belonging to different dimensions—which the mind, through *memory*, recognizes as being associated in some way or deriving from the same source—represent the *description* of an "object." As intelligent human cognitive systems, when we observe "the pen on the desk," "the neighbor's house," "the just-purchased candy," "our aunt," we have no doubt we are "observing" an "object"—possessing color, weight, hardness, flavor—even if we in fact perceive "sensorial stimuli" (visual, tactile, auditory, etc.) which our mind "associates" in a unitary system: "the pen," "the house," "the candy" and "the person." I have written "the pen on the desk" in quotes to indicate reference to a particular observed item; that is, a particular pen. The need to use language to describe the observed operations entails the need as well to distinguish between "words" to indicate the observed objects and the concepts (Sect. 9.6).

Nevertheless, we are accustomed to also considering as "objects" a window, a hand, a road, an intersection, even if it is difficult to perceive the weight of a window, the height of a hand, the flavor of a road, or the color of an intersection.

The mind carries out the same basic activities for "objects" that it does with perceptions: it perceives the objects, already present at an unconscious level, and compares them to uncover differences or equalities. This comparison among objects and the search for differences or equalities among objects represents the *base mental operations* of conscious cognitive activity. The mind forms at an unconscious level the observed dimensions (with any eventual dimensional distances) as well as the objects of observation. The dimensions and the objects represent the fundamental elements of conscious observation for the mind.

Using a simple formal notation, in order to operationally translate Bateson's difference-based epistemology, "describing" an object "A" means:

(a) Identifying or choosing a convenient number N of *dimensions*—or *scales of differences*—$D_1, D_2, \ldots, D_n, \ldots, D_N$, that denominate *observable dimensions*, which depend on the structure of the "mind-processor." The ordered set of observable dimensions forms the *observed universe* (UN) adopted by the mind for its descriptions:

$$UN = [D_1, D_2, \ldots, D_n, \ldots, D_N]; \qquad (9.1)$$

(b) Identifying the differences observed in (associated with) object "A"—and which constitute "A"—by specifying the state, $d_n(A)$, that each D_n assumes in object "A," obtained through a precise process of *qualitative* or *quantitative determination* (according to the limits of the "mind-processor"). The vector of the determinations (differences) $d_n(A)$ for each $D_n \subset UN$ forms the *technical description* of "A," which I shall represent with the following notation:

$$[\text{des A}] = [d_1(A), d_2(A), \ldots, d_n(A), \ldots, d_N(A)] \qquad (9.2)$$

We must assume that each object "A"—for the "mind-processor" of a given cognitive system—corresponds to a *description* assuming the form in (9.2) achieved through the differentiation process indicated in (9.1). We can apply (9.2) for the technical descriptions of some of the objects we have encountered above:

boy b o y Bo**Y** B*O*b**O** <u>Bo**A**</u> <u>BOa</u> <u>B</u> **O a** b o Å **A a** <u>B</u> *B* <u>a</u>

Let us assume the sensors allow us to construct the following UN:

UN = [D_1 = "letter," D_2 = "up/low case," D_3 = "color," D_4 = "font," D_5 = " size," D_6 = "distance from next object," D_7 = "is part of a word"]

The descriptions of some of the objects are represented by the following vectors:

[des **b**] = [b, lowercase, black, times, 10 pt., none, "is" part of a word]
[des **b**] = [b, lowercase, black, Arial black, 10 pt., 3 pt., "may be part" of a word]
[des <u>B</u>] = [b, uppercase, red, times, 10 pt., none, "may be" part of a word]
[des **a**] = [a, lowercase, red, Chicago, 16 pt., 3 pt., is not part of a word], etc.

The *definition* process is always *relative*, since it depends on the *number* of dimensions considered in UN and the *precision* of the specification of their state, which the "mind" can produce through some process of qualitative or quantitative *determination*.

The *descriptions* can be:

- *Punctual,* if they concern a single object, or *general*, if they concern the dimensions common to the objects of a given set.
- *Static,* if the variables are observed at an instant in time, or independent of the time.

- Variable, or *dynamic*, if they also concern the variation in the state of the variables with respect to another variable that the "mind" conceives as "time"; the *dynamic descriptions* assume the search for differences among the differences observed at different moments, which in turn assumes the use of memory.

The *dynamic punctual technical description*, with reference to period T, can assume the following structure:

$$[\text{des } A(t)] = [d_1(A,t), d_2(A,t), \ldots, d_n(A,t), \ldots, d_N(A,t)], t \text{ in } T. \qquad (9.3)$$

I have defined (9.2) and (9.3) as *technical descriptions* since they are independent of any linguistic representation; they are the result of a "mental process." However, they can be translated into *linguistic* descriptions using a chosen language (Sect. 9.10).

Without entering into overly complex cognitive problems, but only in order to justify Bateson's idea of mind, I would note that we humans are cognitive beings possessing a very powerful and versatile mind that allows us to identify, with incredible speed, such a large number of types of differences that we seem to be directly perceiving objects as if they were "sources" of those differences. As a result, for humans [des A] seems to derive directly from object "A." However, this is true only for objects we can perceive by sight. Try to identify an object in the dark using the many differences perceived through touch; or to identify food by its taste with your eyes covered, or aroma and the many differences perceived with the sensors on the tongue. You will realize the object is not the source of your perceptions but instead derives from these, since the mind "associates" them by forming a perceptual unit. Some further considerations will clarify the *human* "potency" of the mental operation of constructing *technical descriptions*, even if what we have learned during our lifetimes and through the use of language oblige us to make an effort not to think directly of the objects but to concentrate on their dimensions, D_n, and their states, d_n.

Above all, I would observe that the human mind is predisposed to perceiving (detecting) numerous types of *differences* and to ordering these into an equally numerous amount of *dimensions*. In fact, we are used to defining *dimension* (connotation, characteristic, parameter, or way an object manifests itself) as any *association of differences* judged to be similar, which the mind can *perceive* in defined and given spatial-temporal contexts and which, as humans, we associate with "objects." Operationally speaking, the *dimensions* can thus be viewed as *generalized modes of perception* (weight, color, flavor, etc.) or of *association* (name, relative position, etc.) regarding certain *differences* (perceivable, defined, or imaginary) that a given observer judges to belong to the same species even though he places them in different objects. These dimensions are not "observable" but are "constructed" by the mind in order to compare various objects in relation to the states taken on by the dimensions created for purposes of comparison. Any further calculation of differences regarding the objects absolutely requires that the mind be able to determine in some way the state taken on by the dimensions "in" the various objects.
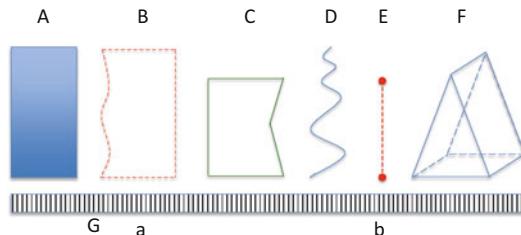
Fig. 9.3 Differences and dimensions

I shall define as *determination* the mental operation by which the mind, in making comparisons or employing first-level dimensional states which have already been *memorized*, specifies the state of each dimension of the technical description of an object in a given observable universe. To make this aspect even clearer, let us look at the example in Fig. 9.3, which, unfortunately, must necessarily refer to the human mind.

Through the operation of sensorial (sight, touch, etc.) and mental (memory, concentration) "organs," based on given modalities which we shall not examine here, the mind is able to recognize Fig. 9.3 in its "entirety" as a *unitary object* through the perception and arrangement of the differences found (traces of ink of various colors) in comparison to the uniformity present (blank white sheet of paper). In fact, according to Bateson's model, the mind, as a determiner of differences, perceives the black traces on the white background rather than the white background *sections* that remain after considering the black spots (the possibilities afforded us by our capacity to continually observe a particular object rather than its complement is self-evident, which is brought out even more when we observe *ambivalent or ambiguous figures*). The mind can therefore consider the *unitary object* (that is, the overall figure) as a composite object (Sect. 9.5) and *isolate* in it a variety of more *circumscribed* objects; it can focus attention on the hatched box at the bottom, which, observing its relative (difference) position can be interpreted as a "plane of reference." The other figures "rising up" from this plane of reference can then be observed, in addition to the symbols we recognize as letters of the alphabet.

Concentrating our observation on the figures "rising up" from the base, the mind can perceive differences in each with respect to the others; for example, all are linked to other figures (which we recognize as capital letters) placed above them, and each figure is "erected" differently with respect to the plane of reference and "rests" on it in a different manner. We can thus clearly grasp the process involving the "construction" of the *dimensions*, understood as generalizations of the perception of differences which are distinct yet held to be of the same type. The perception, for example, that all the figures are linked to a particular figure, a different letter of the alphabet, can lead the mind to generalize the perception of these differences (letters) as distinct "states" of the dimension "designation" (*acronym*, *denomination*, etc.). Through extrapolation the mind generalizes that the "name" represents a characteristic common to all the figures, and that for each of these this dimension presents a particular

"state"; in the first figure on the left (according to our human perspective), the dimension takes on "state" A, in the first on the right, "state" F, and so on.

Perceiving in all the figures a different "*upward*" movement with respect to the designation G (plane of reference) leads us to generalize the "*height dimension*" of those figures; perceiving a difference in the area on which the figures rest on the plane of reference leads to the generalization "*width dimension*"; through association with other types of differences the mind perceives in all the specific figures, which it considers to be homogeneous, the mind makes other generalizations regarding the dimensions, for example, "color," "shape," "perimeter," "surface" (total, lateral, base), "position," and so on, according to the mind's ability to find similarities among the dimensional states perceived in the particular figures. We could also state (using our mind, which for years now has already been accustomed to carrying out these operations) that the creation of dimensions also depends on the ability of the cognitive subject and on the instruments he adopts and the objectives he pursues.

To conclude this example, I would maintain that in Bateson's model the *dimensions* are created through the mental process involving the *generalization of* specific *differences* perceived in the observed objects, which are considered similar with respect to a common and unvarying *point of reference*. Thus, this example allows us to conclude that the relation [states → dimensions → objects] can be inverted: [objects → states → dimensions], without getting away from Bateson's logic. I do not think the idea would have "entered into the mind" of any human observer (with vast experience) of also associating with the specific objects in Fig. 9.3 the dimensions of *weight, specific weight,* or *brand of computer with which the figures were traced*. In fact, in Fig. 9.3 none of these dimensions manifest particular "perceivable states" that, through generalization, allow one to create those dimensions. Perhaps only the figure "designated" as F, which is observed as being three-dimensional, could have a "state" denominated as "weight" and one as "volume." The figures "denominated" D and E manifest the dimension "length," but not that of "surface," which instead characterizes the figures denominated A, B, and C. On the other hand, a mind could attribute to the figure "denominated" G the dimension "length," as proof that the dimensions are constructed by the mind through the generalization of quantifiable specific differences in the objects. Until now I have ignored the two signs "a" and "b" corresponding to the figures denominated "B" and "E," below the figure denominated "G." In which dimensions can they be compared? Perhaps the mind, with the obvious dimension *capital letter/lowercase letter* excluded, could consider "a" and "b" as "denominations" for only the two objects whose state is "dotted" in the *type of line* dimension. However, I do not exclude the possibility some minds might compare these using other dimensions.

The above example shows, even if the distinction does not possess absolute validity, the possibility of considering separately the *observed* dimensions with respect to those *associated* with a particular object of observation. The dimensions *observed* represent differences the human observer feels can be detected in the object: color, weight, volume, function, shape, flavor, material of construction, composition, structure, observation date, etc. The dimensions *associated*

with objects, on the other hand, are determined by the observer himself, if necessary by developing dimensions from another category. The relative position of the object with respect to the observer, the name, specific weight, average weight, prevalent color, code number, name of builder, date of construction, weight per square decimeter of the base, and so on, represent examples of associable dimensions. Since the distinction is relative it is not always easy to allocate the dimensions to one or the other category. If we consider that all observations are extremely subjective, then all the dimensions are relative to the mind of the observer, and for this reason can be considered *associated* with the object. If, on the other hand, we consider the object as a unit that is directly describable, then not only would the dimensions held to "naturally" emerge from the object's existence be *detectable* but so too would extension, mass, electrical charge, as well as any difference inherent in the object, without which the object "would not be" observable. Thus the *weight* dimension can be considered one which is *observable* in three-dimensional material objects (every object has— must have—a weight) or capable of being *associable* for other objects (every object has a weight only if observed in an environment with gravity). An object "has no" weight until a mind produces it by calculating a difference within a given observational universe (9.1).

The above considerations imply that, at least *in man's case*: (1) the technical descriptions of the objects of observation depend on the number, type, and value (state) of the dimensions characterizing them; (2) there is no limit to the number of dimensions by which objects can be described; (3) this number does not depend on the nature of the object but on the number of sensors and the experience (objectives, education, limits) of the observer; (4) we can always imagine that two distinct objects, P and Q, can be described in the same *observed universe*, even if some dimensions can be used only for P and not for Q; we only need consider the dimensions that do not characterize Q as dimensions with a "null state," for which no difference in Q can be perceived; (5) to make the technical descriptions operationally efficient, we must *limit* the number of dimensions taken into account for the actual [des A] in (9.2).

9.5   Comparing Objects

When the "mind" applies the process for recognizing the *differences of differences* to the technical descriptions, it is then able to distinguish between objects that are *equal* or *different*.

I define the *equality* of "objects" by writing:

$$A = B \text{ in [UN] if, for each } D_n \subset \text{[UN], we have} : d_n(A) = d_n(B); \tag{9.4}$$
$$\text{that is, if } [\text{des } A = \text{des } B].$$

Fig. 9.4 A comparison of objects through their technical description

I define the *difference* of objects by writing:

$$A \neq B \text{ in [UN] if, } \textit{for at least one } D_n, \text{ we have } d_n(A) \neq d_n(B);$$
$$\text{that is, if } [\text{des } A \neq \text{des } B]$$

(9.5)

Comparing A and B through their technical descriptions, [des A] and [des B], assumes the use of *Rings* operating over many cycles and at high frequency, similar to those shown in Fig. 9.4.

It is important to observe that *equality* or *difference* in objects can be affirmed only after we have defined the *observable universe*, [UN]. Many errors in observation occur because a [UN] of reference is not specified. Presumably we would say that two coins which have the same "face value," same shape, same weight, same diameter, same thickness, and are made of the same metal are equal. If we widen the observable universe to include the dimensions "year of coinage," "owner," "number of component molecules," etc., then it becomes difficult to perceive *equality* in those objects. If we also included in the observable universe the dimension "spatial collocation" (the same spatial coordinates), then the two objects surely cannot be *equal* in the sense specified above.

Comparing the *technical descriptions* of objects also leads to distinguishing between *simple and composite objects*. An object is *simple* if, for the "mind" and/or for a specific observation, it is held to by *unitary*; that is, it cannot be broken up into other elementary objects, B, C, etc., so that there is no sense in holding that

[des A] can be broken up into autonomous [des B], [des C], etc. An object A is *composite* if it can be considered to be composed of (or broken up into) autonomous components (parts), B, C, etc., which can be independently described; in other words, if it makes sense to obtain [des A] = [des B] + [des C].

The mind also distinguishes between *separate* and *united* objects. The objects A and B are *separate* if, *for each* $D_n \subset$ [UN], [des A] can be formed through determinations which are independent of those necessary to form [des B]. If there is a D_n such that we can observe only $d_n(A\&B)$, and it is not possible to observe $d_n(A)$ separately from $d_n(B)$, then the objects A and B are *united* in (for) that dimension. Arm, hand, and fingers are *united* objects but hand, wedding ring, and watch are *separate* objects.

Two objects, A&B, though *united* in terms of the values of the dimension D_n, can be *separable* if, through some *conventional procedure*, we can determine the values $d_n(A)$ and $d_n(B)$ and include these separately in [des A] and [des B]. If there is at least one D_n which does not admit any conventional procedure to independently determine the states $d_n(A)$ and $d_n(B)$, then the A&B are *not separable* in the absolute sense. A four-room apartment, for example, can be broken up with respect to surface area, wall color, and contents into four objects corresponding to the four individually describable rooms; nevertheless, the apartment can be sold while the individual rooms cannot. With respect to the selling price, the apartment is a nonseparable object. Even a segment can be broken up in its description into four segments, with respect to length, even if there is still only a single segment. If a pair of oxen are pulling a cart it is not possible to observe as two objects the individual oxen each pulling half a cart. Two *separable* objects, A and B, can be *united* to form the object A&B if it is possible to form [des A&B] so that for each "*n*": $d_n(A \& B) = d_n(A) + d_n(B)$.

Two *united* objects, A and B, are *autonomous* if they can be described by different dimensions, leaving out only those dimensions for which they prove to be united. A finger is an object we all recognize as *autonomous* with respect to the palm of the hand, even if it is *united* to it, because we believe the *finger* and *palm* to be different in structure and function and consider it irrelevant to also include in the two *technical descriptions* the color and weight of both the finger and palm. A tree trunk can be independently described with respect to its foliage and roots, even if it is united to these.

The objects which can only be *described* through *differences* detected with the sensors (detected states of dimensions) can be defined as *real objects* for the observer. Referring to the "human mind," the use of our *associative capacities* (Sect. 9.2) often leads us to include in [des O] dimensional states which are not observable or not pertinent, thereby creating imaginary objects. I shall define as *imaginary* those "objects" deriving from the mental activity that, using new terms or based on unusual forms, "associates" values (differences) of dimensions *not observable* in those objects but deriving from the observation of other objects, as occurs, for example, when we associate a color to an electron, describe a three-headed dog, a unicorn, a winged-steed, or a flying carpet. This human capacity to associate nonobservable dimensional states, including these in the *description of*

objects, represents the "imagination" that can also lead even to conceiving of highly complex imaginary objects, which we are all familiar with from having read the many imaginary descriptions that abound not only in science fiction and fiction books but also in many classical works, for example Dante's Divine Comedy, John Milton's Paradise Lost, etc. As they cannot be observed, *imaginary objects*, though they can be *enunciated*, can also be called *nominal*. Some of the *imaginary objects* which observation has not yet shown to exist, but neither demonstrated do not exist, can be defined as *hypothetical* (for example, antimatter, mathematical "strings," soul, etc.).

9.6   The Second Step Toward Knowledge: The Process of *Definition*: From Objects to Concepts

Mental activity is incessant. On the one hand, the mind perceives *differences*, applying the calculation of these differences to the differences themselves (different levels of differences) and "associates" differences by forming technical descriptions of "objects." On the other hand, the mind compares the technical descriptions and uncovers *analogies*. If identifying differences is necessary in order to perceive "objects," analogy is indispensable for *generalizing differences* and constructing "dimensions" and associating into categories objects which, while different, are held to be "similar" or "analogous" even though their technical descriptions are subject to change.

Technical descriptions represent the first step in acquiring knowledge. The second step is that by which, through *analogy*, the "mind" transforms the "technical descriptions" of objects into "technical definitions" from which we form *concepts* (ideas) of objects. A "technical definition" is a *"general" description* that specifies the *limits* of variability in the "technical descriptions" of *objects* in order that these can be considered *analogous;* that is, all belonging to the same *open set* represented by the definition (genus, species, type, class, etc.).

The *process of defining concepts* thus implies an *analogical generalization* through which the differences found in various "technical descriptions" of objects are evaluated by the "mind" (calculating differences of differences) in order to "conserve" the *analogies* and to form a *general class of objects* that we can then associate with a *generic object*.

In formal terms, if, for each $D_n \subset UN$, the symbol $\Delta d_n(O^*)$ indicates *the range of admissible variation* of the dimensional states in order for an *object* "O" to be part of the *concept* "O*" (idea, intuition, general significance, abstraction, generalization, image, etc.), then the "technical definition" of "O*" is represented by the vector:

$$[\text{def } O^*] = [\Delta d_1(O^*), \Delta d_2(O^*), \ldots, \Delta d_n(O^*), \ldots, \Delta d_N(O^*)] \subset U(N) \quad (9.6)$$
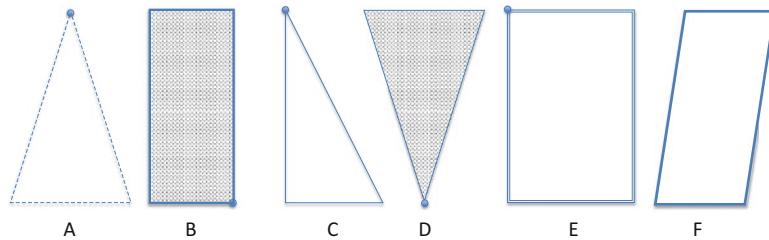
Fig. 9.5 Producing a technical definition

Each "technical description" can thus be considered a "particular case," an "example," of a "technical definition." I can then state that for the knowing "mind" the *object* "O" is part of the *concept* "O*" if [des O] ⊆ [def O*]. More simply put, each specific object can be considered an "example" of the "idea" of a "generic object." I describe "*that*" particular chair but I define "chair"; "*that*" chair is an example of "chair."

The *definition* can be formulated using two processes:

1. *Denotative* or *extensive*, if the process identifies the set (denotation or extension) of the descriptions [des O_1], [des O_2], [des O_3], etc., that can be included in [def O*].
2. *Connotative* or *intensive,* if it identifies the *dimension* and *range of variation* within which the *states* $d_n(O)$ identified in the objects (in their descriptions) must be included in order for them to be indicated by the definition (connotative or intensive). The connotative definition is provided in (9.6).

Thus, for example, all the objects denominated A to F in Fig. 9.5 are undoubtedly different from one another, given the technical descriptions the reader can make using (9.2) and considering the following simplified UN:

UN = [D_1 = "number of sides," D_2 = "shape of sides," D_3 = "length of sides," D_4 = "vastness of area," D_5 = "surface color," D_6 = "direction of vertex (point)," D_7 = "name," D_7 = "position," etc.]

The generalizing activity could hold, for example, that the figures A, C, and D are *analogous*, since they have the same "number of sides," independently of their "shape," "height," "breadth," "width" and "surface color." Figures B, E, and F could also be considered *analogous* in terms of "number of sides." However, another mind might have considered as *analogous* those figures whose sides are a certain thickness, independently of the number of sides, while a third mind could have viewed as *analogous* those figures whose vertices (indicated by a solid circle) are pointing north. Analogical activity, being mental, is always subjective.

We can generalize by going back to the example in Fig. 9.5 and keeping (9.6) in mind. Clearly two or more objects of observation are held to be "analogous" if they are, above all, characterized by the same UN and, secondly, if the differences in the quantity and quality of the individual d_n included in the *technical description* fall within the *ranges of admissible variation* defined on each occasion by the mind in order for there to exist a similarity.

This means that two objects, A^I and A^J, are considered *analogous* if, though their technical descriptions differ, the individual differences in quality and quantity are held to be included within the *range of admissible variation* characterizing "the" object A* in general: that is, the concept A* (for the moment we have to resist the urge to use language to denominate A*). The mind can now *generalize* as follows: the objects A, C, and D from the previous example, whose technical descriptions differ, are considered *analogous* because the states (differences) of the variables (scales of differences) from their technical descriptions (association of differences) revealed by the sensors are not "too different" from one another; using *analogical generalization*, the three objects can be thought of as special cases of a general object: "triangle," having the following technical definition (referring always to the shapes in Fig. 9.5):

[def TRIANGLE*] =
=["number of sides" = three and only three (no range of variation)
    "shape of sides" = any whatsoever (infinite range of variation)
    "length of sides" = any whatsoever, as long as it is positive (unless we are considering figures with negative measurements for the sides with respect to the origin) (infinite range of variation)
    "width of area": correlated with the length of the sides
    "surface color": any whatsoever
    "direction of highlighted vertex": any whatsoever].

We can reverse the process: given the *technical definition* of the general object A*, how can we verify if the specific objects, A^I and A^J, for which the mind has calculated [des A^I] and [des A^J], are similar? The answer is simple: we need to verify that each *value* for the technical descriptions is included in the *admissible range of variation for the existence of the analogy*. Only if the values for all the dimensions, $d_n(A^I)$ and $d_n(A^J)$, of the two objects are included in $\Delta d_n(A^*)$ are the objects *analogous* and can be considered as "examples" ("→" means "example of") of the concept [def A*]. Therefore, indicating analogy with the symbol "≈," we can write the result of the comparison of the similar objects in one of the following ways (Fig. 9.6):

$$A^I \approx A^J, \text{that is } \left[A^I, A^J\right] \rightarrow A^*$$
$$\text{if } \left[\text{des } A^I\right] \approx \left[\text{des } A^J\right], \text{that is}$$
$$\left\{ \left[\text{des } A^I\right], \left[\text{des } A^J\right] \right\} \subset \left[\text{def } A^*\right], \text{that is}$$
$$d_n\left(A^I\right) \subset \Delta d_n(A^*) \; e \; d_n\left(A^J\right) \subset \Delta d_n(A^*), \text{for each "}n\text{."}$$

With the aid of the above example, we can immediately deduce that the general object "A* = "triangle" derives from the *technical definition* "closed three-sided figure (figure with three angles)," which is easily obtainable from the preceding considerations and as a result capable of being formalized. Figures A, C, and D represent *analogous* objects (that is, they are specific triangles) since their technical description is included in the technical definition of the *general object* "triangle."
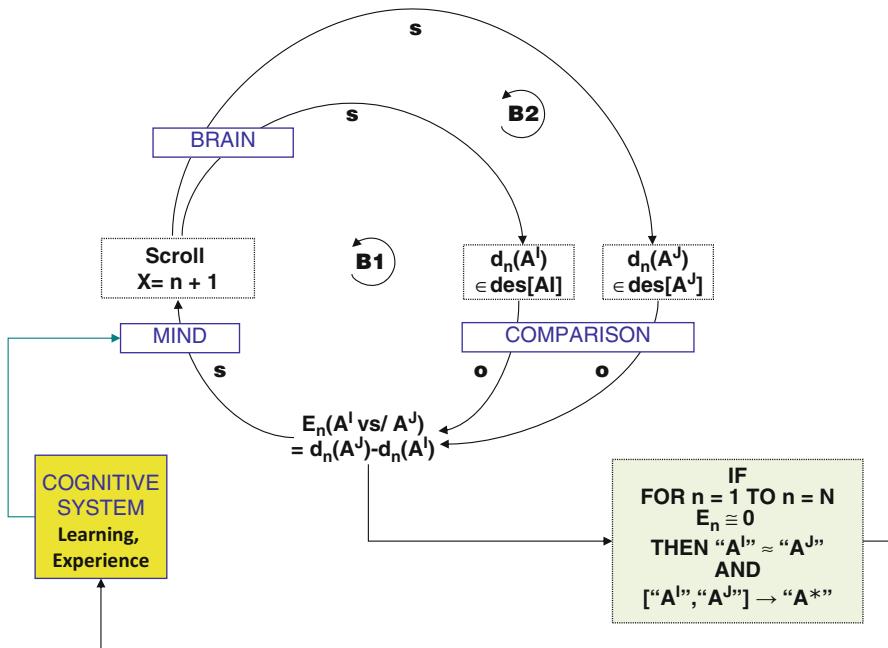
Fig. 9.6 Verifying the similarity of the objects and their belonging to a "concept"

The *general object* "the triangle" is a *conceptual abstraction* connected to the *analogical* and *generalizing* activities of the mind. In fact, the *general object* represents the set (defined intensively) of all specific objects (the individual triangles that can be observed or imagined) for which an analogy exists in the mind of a given cognitive system; that is, the set of specific objects whose technical *description* is included in the technical *definition* of the general object.

Objects are described; *concepts* are defined.

I shall define "a thing" as that which the mind does not define; this is a concept that includes every describable object. The technical definition of "a thing" is so general as to include as particular cases any specific object of observation. Formally, we can write:

$$[\text{def THING}^*] = [\Delta d_i(\text{THING}^*) = \text{"whatever, "for each "}i\text{"}]$$

This concept is in line with the use we make of the concept of "thing" when (through language) we want to designate unrecognized or unrecognizable objects of observation or to indicate any object or concept that is not described or defined, for example: "Some*thing* strange was seen in the sky"; "There are many *things* on the table; bring me all of them"; "The specific weight can be determined for all *things* that have a volume and a weight"; "All the visible and invisible *things* have been created by God"; etc.

9.7   The Three Paths of Knowledge in a Two-Dimensional Cognitive Universe

Following Bateson's assumptions, we can now verify how the mind-processor—by means of the four basic mental operations *distinction, comparison, perception of analogies* and association (Sect. 9.2)—can gain *knowledge* by developing *three minimal cognitive processes* at both the conscious and unconscious levels (Fig. 9.7):

1. *Cognition, or perception, of objects*: adopting the mental processes of *perception* and *comparison,* and employing constitutive memory, the mind forms *descriptions* of objects—[des A], [des B], [des C], etc.—distinguishing these from their *background*.
2. *Understanding*: after having described several objects, the mind uses *analogical abstraction* to develop [def A*], [def B*], [def C*], etc.; that is, it finds in the *observational universe* a *zone* that corresponds to [def A*], a second zone corresponding to [def B*], a third zone corresponding to [def C*], and so on.
3. *Recognition*: in the presence of a specific object "O," the mind carries out the [des O] and determines in which [def O*] [des O] is included. The object "O"
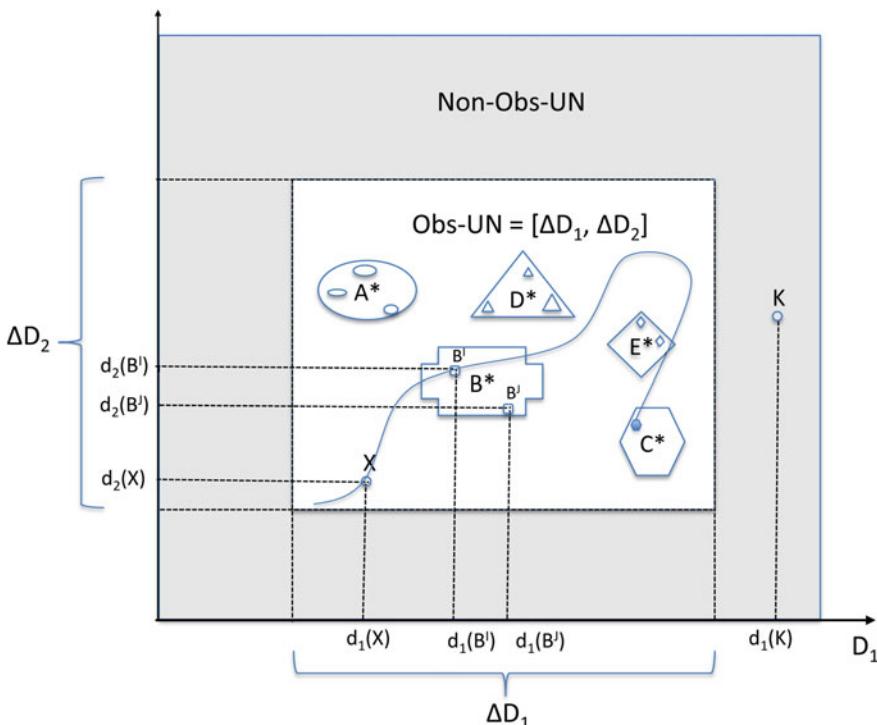
Fig. 9.7 Observable universe, descriptions and definitions

then becomes a *point* in [def O*]. If the *recognition* is successful, the mind *recognizes* (*knows*) "object O" as belonging to "concept O*." In the opposite case it determines a new [def O*].

Referring to our experience, we can easily realize that the cognitive processes are the basis for understanding, allowing us to form the [def O*] of various objects through an analogical process applied—at times unconsciously—to various technical descriptions; often the initial [def O*] coincides with a simple [des O]. "Knowledge" is acquired through a continuous cycle of "*cognition*," "*understanding*" and "*recognition*," which together form our *minimal thoughts*.

To make the continuous action of the three moments of knowledge clearer, let us consider a "minimum mind" which receives the differences from only sensor organs and thus is only able to distinguish differences belonging to only two dimensions, D_1, D_2, representing a *minimal universe*: UN = $[D_1, D_2]$, shown as the axes in Fig. 9.7. Let us further define ΔD_1 and ΔD_2 as the segments of the two dimensions that form UN, whose values can be distinguished by the sense organs that send the differences to the mind. The area determined by the segments of such dimensions which are knowable to the observer through the available instruments represents the *observable universe*: Obs-UN = $[\Delta D_1, \Delta D_2]$, which, in Fig. 9.7, corresponds to the white rectangle. The grey area represents the universe which cannot be observed by the mind.

The minds of non-evolved biological beings have probably no chance of expanding the *observable universe* beyond the natural limits imposed by their sensory organs and by the range $[\Delta D]$ of differences their minds can perceive. In fact, the nature of the Obs-UN depends fundamentally on the number and quality of sensors made available by the cognitive system to send differences to the mind and allow it to create the dimensions of the universe. It is impossible to imagine a cognitive system without any visual organs forming a technical description of a bunch of colored flowers or a person without the sense of smell forming a technical description of the smell of those flowers. Human minds, on the other hand, are constantly searching for amplifiers for their sensory organs in order to expand the *observable universe* well beyond the limits set by the perceptual ranges of their sensors. Telescopes and microscopes of all types, nocturnal viewers, artificial noses, contact sensors, etc., represent amplifiers that have greatly increased the possibility of perceiving differences, identifying dimensions, and creating increasingly larger composite objects.

Referring to Fig. 9.7, each point of the Obs-UN represents the technical description of an *observable object*. The areas A*, B*, etc., represent *technical definitions*, that is [def A*], [def B*], etc. An object "X" is the point described by the coordinates $d_1(X)$ and $d_2(X)$, which the mind uses to construct [def X] = $[d_1(X), d_2(X)]$. Since the object X does not belong to any of the areas indicated in Obs-UN, the mind can describe X but not understand it, since the concept (the technical definition) to represent it is missing. Completely different is the situation regarding objects B^I and B^J, which can be described by the mind but also understood as objects representing examples of the concept [def B*]. This is represented graphically in Fig. 9.7 by points B^I and B^J included in the area-concept B*.

In general the mind can *know* any object O^H if it describes it in Obs-UN (cognition); it can *understand* it if it detects a similarity with other objects which are "not too different" from it, thus forming the *concept* [def O*] (understanding). When the mind distinguishes another object, O^K, as being similar to O^H, it can *recognize* O^K as belonging to the concept [def O*] (recognition). More generally speaking, any object X *known* (distinguished) by the mind through [des X] is *recognized* if the mind succeeds in including the [des X] in some memorized *technical definition*. It can then recognize that X is A* if [des X] ⊂ [def A*]; X is B* if [des X] ⊂ [def B*], etc. This recognition can fail in two cases:

1. When the mind is not able to (adequately) determine the *technical description* of the object, perceiving "things," not objects.
2. When it is not able to determine the *technical definition* that encompasses that description.

The surface of Obs-UN occupied by the *technical definitions* forms the *known universe*; the set of points that corresponds to objects included in the known universe forms the *observed universe*. If [des X] does not belong to any of the definitions in Obs-UN, then a *new definition* arises represented solely by the point indicating the object: [def X*] = [des X]. Knowledge is increased because the known universe is enriched by the new definition. The curved line in Fig. 9.7 represents an ideal *path of knowledge* of a mind which, through its sensors, explores Obs-UN and, after having perceived X, continues its exploration, perceiving B^I and recognizing this as a point in B*, subsequently moving on to point C*, which it can recognize. I would note, even though it may seem obvious, that the "path" shown in Fig. 9.7 is not in a *spatial* territory but in a *mental* one, which is made up of concepts "mapped" in an *observable universe*.

I would propose a mental experiment which is very difficult for adult humans, since our mind has already constructed all its technical definitions, that is, its concepts, and has associated a specific name to each concept. We must make an effort to simulate a mind that explores a given Obs-UN without already possessing any concepts or a language to name them, but which instead creates the concepts using the technical definitions.

Imagine a baby in a carriage led around by his mother for the first time in the fruit and vegetable section of the supermarket. While the mother pushes the carriage around the various food stands the baby's mind observes, as we can imagine, many differences in color and size that he arranges into "objects." There are many round objects (the baby is still unfamiliar with the concept of "sphere") and different colored ones: yellow, red, green, green and red, red and yellow, etc. (adopting the terms we already know). His mind can recognize similarities and he forms the concept (I must express it in words) of *large round fruit* (which to us is an apple). He then observes many other rounded objects, small but only red, with some variants and with stems: considering them to be similar, he includes them in the concept of *small round fruit with loops* (cherries). Other objects are red and small as well, but without stems and heart-shaped (I use this dimension only for purposes of this example), which he joins together under the concept *red heart-shaped fruit*

(strawberry). Taking the example further, the long yellow objects are placed under the concept *long yellow fruit* (banana); the pear-shaped ones are *pot-bellied pointed fruit*; the dark colored, hairy cylindrical objects are *hairy fruit* (kiwis), etc. The baby is not aware that the objects he observes can be grouped under the concept *fruit*, but I have used this term to make the example easier to understand. The baby returns home and sees on the table a *long yellow fruit*, several *small round fruit* and some *large round fruit*. Soon after the grandmother (concept) arrives with a bag containing some *small round fruit* and some *pot-bellied pointed fruit*.

The above process is repeated week after week, until the baby's mind is able to "denominate" the objects and the "concepts"; that is, he begins to perceive signs-words, form analogies, for example, between the word "pear" pronounced by his mother, father, grandmother, and sister and various *pot-bellied pointed fruit*. He hears "vocal signals" that are different because pronounced by different people that he, nevertheless, judges similar in designating those objects. Finally, one day, again at the supermarket, he points at a *pot-bellied pointed fruit* and his mother says "pear" (pardon my conciseness here). Then he indicates a *large round yellow fruit* and his mother says "apple"; a *large round red fruit*, which the mother again refers to as "apple; a *small round fruit*, which is referred to by the sign "cherry"; the *round heart-shaped* fruit is referred to as "strawberry, while a pointed pot-bellied brown one elicits the sign "pear" from the mother. After a certain number of iterations of the process of perception the baby—thanks to the powerful memory all babies possess—forms his *understanding*; that is, he manages to give a defining content to the technical descriptions of many objects, thereby forming an *observed universe* of concepts to help him recognize various objects of that type.

Once again the grandmother arrives and takes out a *long yellow fruit* from the bag, which the baby recognizes as "banana," saying "*babanana*" as a sign to indicate the concept. His grandmother corrects him: "Not '*babanana*,' "*banana*." The baby recognizes a *small round fruit without a loop* and calls out "cherry." However, the grandmother makes him understand in some way that he has made a mistake. The technical description of that *small round fruit* does not belong to the concept of *cherry* but to that of "cherry tomato." The baby introduces the new concept, thus widening his *observed universe*. The grandmother then takes out a *large round bluish-violet fruit*, causing the baby to exclaim: "apple." Once again the grandmother patiently disrupts the baby's observed universe by saying: "egg-plant," thus producing a new concept.

I shall end the example here, as I feel I have made clear what the *three minimal cognitive processes* are while also generalizing them. We establish the observed universe in which the *mind* perceives differences and objects by comparing them through a set of dimensions it constructs; if in the chosen Obs-UN the mind *perceives* "n" different objects and by means of the dimensions forms the [des O^n], which it holds to be *analogous*, then the mind *understands* by forming a *concept* by producing the [def O^n]. If the *mind perceives a new object* through [des O^x], it *recognizes* that object if it can identify the [def $O*$] that represents the concept, using the same *Ring* of *recognition* we know from Fig. 3.3, which we can now adapt as shown in Fig. 9.8.
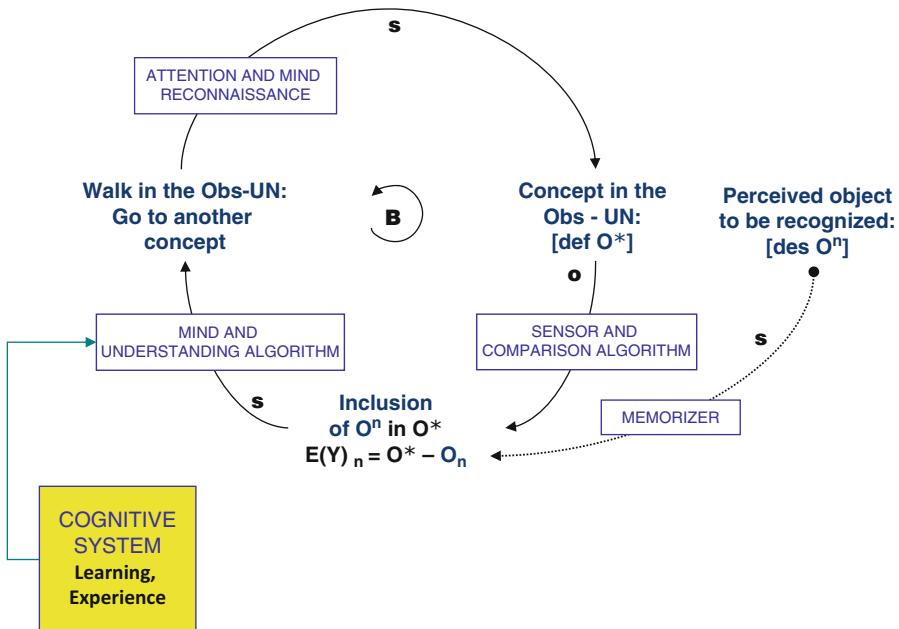
Fig. 9.8 The *Ring* for recognizing-identifying objects

The importance of the *observed* universe is clear. Changing the UN means that the knowledge processes must also be restructured. To show this, let us continue with our example and assume that one day the baby *tastes* a slice of apple. The grandmother asks him: "What fruit is that?" The baby cannot avoid becoming aware of the new perception in the "flavor" dimension, but obviously he cannot recognize the fruit since its flavor was not a dimension he used to describe different fruits and to recognize them by means of their respective concepts. Once he has tried several times a sample of all the known fruits (whose definition he possesses) and associated the new states of the "flavor" dimension to the perceived objects, his knowledge is broadened and he is able to recognize a particular fruit either by looking at it or by tasting it. In fact, before tasting them the baby observed the fruits in the three-dimensional universe:

$$UN = [D_1 = \text{“shape,”} D_2 = \text{“dimension,”} D_3 = \text{“color”}].$$

After having distinguished the flavors and with the "flavor" dimension thus added, the universe of observation becomes four-dimensional:

$$UN = [D_1 = \text{“shape,”} D_2 = \text{“dimension,”} D_3 = \text{“color,”} D_4 = \text{“flavor”}]$$

This broadening of the UN noticeably increases the capacity to form concepts and, above all, to carry out recognitions.

What happens if the mind perceives a new "object" not part of an acquired "concept"? As an example, assume we want to scrutinize the sky using modern technology and to distinguish "something new," with "X" indicating the "new object" that has to be recognized. How does the cognitive process unfold? The first step is to compose the [des X] using the dimensions that normally characterize well-known celestial objects. Subsequently we search for a *technical definition*, already set forth, that allows us to *recognize* the object X as analogous to others already observed. If we are successful in our search, knowledge is acquired and we can conclude: "since object X has dimensional states that we already know, we can conclude, for example, that it is a *pulsar*." If our search for the *technical definition* is not successful, in the sense that "X" continues to remain a "mysterious object," we can create a new *technical definition* and conclude: "in that zone of the sky a *new* celestial object was observed for the first time that presents the dimensional states [des X] = [$d_1(X), d_2(X), \ldots, d_n(X), \ldots$]; since it is not analogous to any of the celestial objects already defined (known), the new celestial object (and any others that have similar characteristics or dimensional states) will be *called*, for example, "MP13." The new *technical definition* (which in this case coincides with only a single technical description) has added an additional element to the *defined universe* and raised our capacity for understanding and recognition celestial "objects."

9.8   The Human Mind Constructs Meaningful Technical Definitions

The mind would not be efficient in operating as a calculator of differences if, in addition to the activity of *perception* and *analogy*, it did not also carry out *simplifying* activities. The continual search for "simplicity" (even though this term is subjective) pushes man toward "simplification," without which any generalization would be become extremely challenging in terms of efficient observation. Analogy is efficient only through simplified perceptions, descriptions, and definitions, since the mind tends to ignore anything appearing "superfluous" so as to concentrate on the perception of dimensions thought to be "relevant" for knowledge. Thus, above all in formulating descriptions or definitions, the minds of human beings tend to ignore many dimensions of the observable universe, since their inclusion would make the descriptive and definition process so complex and redundant as to render it inefficient. For this reason there is an attempt to reduce to a minimum the number of dimensions considered in the descriptive and defining activities in the knowledge process in order to form *minimal* descriptions and definitions, which we shall call "*meaningful descriptions and definitions*" (or *basic* concepts): [min des O] and [min def O*].

Many *meaningful descriptions* and *definitions* even include a single dimension held to be particularly representative; these are called *elementary*. Below are some

of these types of definitions, expressed, for simplicity's sake, in a linguistic form (the typology is also valid for descriptions):

(a) *Ostensive* definition: this lets the observer form the analogy by merely indicating several examples of the object of the analogical abstraction ("what is an ant?"; "any insect similar to those you see on the tree!").
(b) *Extensive* definition: this lists "all" the objects (extension) that must be contained in the definition ("what is a sisar?"; "any stellar object whose definition is listed on p. 22 of the Atlantis of the Sky").
(c) *Genetic* definition: this highlights the origins of the objects to be included in the definition ("an American is any individual born in the U.S.").
(d) *Historical–geographic definition*: this considers the place and time the objects can be observed.
(e) *Structural* definition: this brings out the structure of the defined objects ("a hand is a limb composed of the following elements ...").
(f) *Functional* definition: this brings out the function of the objects in the definition;
(g) *Modal* definition: this indicates the composition of the objects contained in the definition (materials, color, shape, etc.).
(h) *Instrumental* definition: this brings to light the possibilities for using the defined objects.
(i) *Teleological* definition: this considers the objectives of the objects of observation.
(j) *Operational* definition: this specifies the operations needed to identify or recognize the objects of observation in the definition. Operational definitions are particularly effective in defining abstract or composite objects.
(k) *Stipulative* definition: a definition created ad hoc for a specific purpose; for example, a specific communication or argumentation.

The meaningful definitions commonly used in human knowledge derived from combinations of a limited number of the preceding types. Some further considerations will clarify the different types of meaningful technical definitions.

Ostensive definitions are not true definitions, since they do not, in fact, highlight a specific dimension. Through ostension we can merely indicate one or a few objects which exemplify the definition we wish to form, leaving it to the cognitive subject to apply the analogy to the technical descriptions of the objects by identifying the meaningful dimensions. Just as in the example with the baby, ostensive definitions are usually formed as an answer to a question. To the question: "What is a pineapple?" the mother indicates several pineapples on the supermarket stand, letting the baby make the generalization "large rounded fruit with protruding needles," or something similar. As can be deduced, *ostensive definitions* are generally very meaningful, but only for the mind that has created them; they are not easily generalizable since the dimensions chosen are completely subjective.

Genetic definitions do not directly consider the dimensions of shape, weight, color, etc., but instead focus on the moments, causes, factors, and procedures that have given rise to the objects observed. Thus, for example, we could define "bread"

as the product obtained from a baker; "pizza" as the product eaten in pizzerias, etc. The genetic definition of a prime number could refer to the procedure for verifying this: "a prime number is any whole number divisible only by 1 or itself." These definitions can be called genetic since they illustrate the way the objects that can be included in the definitions are produced.

Historical–geographic definitions are those that point out the historical or geographical location of the objects that must be included in them. For example, "Eskimos" are "individuals from the extreme north of Alaska, Canada and Greenland." "Crocodiles" are "reptiles that live in rivers, lakes and swamps of Central and South America, mainly in the tropical band." Obviously such definitions are not always meaningful, and for this reason they are associated with other types of definitions.

Much more significant are the *structural* definitions that present the dimensions consisting of the "structural conformation" of the objects of observation; that is, their system of components, parts, organs, essential elements, etc. together with their "organization." The "arm" is "each of the upper limbs of a human, composed of two parts joined by the elbow; one part begins at the shoulder and the other ends at the wrist." This structural definition is somewhat approximate though very significant, since it allows the observer to immediately identify the examples after having made the technical description.

If we wish to extend the definition of "arm" to that of "robot" as well, we can construct the *functional definitions*, which indicate several possible and evident ways the objects of observation function. "Arm" is "a structure, biological or mechanical, with a long and relatively thin shape that extends from a trunk and that, through two articulations joined by an elbow, permits a prehensile mechanism (a hand, a claw, an artificial limb) to move." This definition is very meaningful and general, since it also includes the biological structures of animals or of other mechanisms.

I shall define *modal definitions* as those that include the visual or tactile dimension of the objects or elements in the definition or those characteristics that highlight the differing fundamental modalities. "Lemon" is any yellow fruit with a rind and a sour taste." "Muscat grape wine" is "any sweet, fizzy dessert wine."

The *instrumental definition* is also very significant. This definition is applied to objects of observation instrumental to a particular objective. In fact, this definition brings out the different admissible alternatives, procedures, modalities, etc., regarding the use of the objects. A "chair" is "an instrument used to sit down." This definition does not refer to the place where the objects can be observed (homes, cafes, schools, etc.), their structure (with or without wheels, with an inclinable or rigid back, etc.), or the material they are made of (wood, steel, Plexiglas, etc.), only to their use. In many cases these definitions are sufficiently significant, in others they must be aided by recourse to other dimensions, for example, structure (the seat must not have arms, otherwise it is an armchair; it must have a backrest, otherwise it is a stool) and height (a chair that is too low is a footstool). It is not always easy to distinguish between *functional definitions* and *instrumental definitions*, since the function of both can be linked to what the object of observation can be used for.

   Objects of observation can be systems of various kinds which the mind considers characterized by independent objectives that are achieved by means of teleological behavior. Often objectives are attributed to the objects. *Teleological* definitions are those that aim at emphasizing the objectives that characterize or are attributed to the objects of observation. For example, we can define "enterprises" as organizations existing in a market economy and whose ultimate objective is to create profit and stockholder value through the production or exchange of goods or services.

   Meaningful definitions, though they allow us to identify the objects of observation, often do not place the observer in a position to actually search for these objects. *Operational definitions* are those that include elements—procedures, heurisms, algorithms, etc.—that permit the observer to search for and identify the objects of observation. When is a number a "perfect square?"; when this number is obtained by multiplying a whole number by itself." What is a "physical field?"; roughly speaking, it is "the set of values for each point in space of a given physical, scalar or vectoral quantity which is determined through an appropriate procedure and using appropriate instruments."

9.9   Mental Representations of Objects of Observation

The models of [des O] and [def O*] and the concepts of *cognition*, *understanding,* and *recognition* allow us now to translate the symbolic terms we have introduced into "mental" notions that refer to the human mind and of which we make common use. The reader should refer to his own cognitive experiences, which are revealed through behavior and, in most cases, a specific language.

   When a human cognitive subject consciously (by hypothesis) has *cognition* of an object of observation and mentally observes it, his mind probably produces an "image" of that object regarding the spatial and temporal dimensions of the observation. Several *relevant dimensions* are considered which the mind has already formed, guided by attention and memory; others will result from further analyses. However, this initial phase does not lead to any determinations apart from those carried out immediately through simple enumeration (number of sides of a figure, number of legs on a piece of furniture, etc.) or direct observation (color of sides, face of a person, and so on). After the *image* is formed, the observer specifies the dimensions and completes the *cognition* of the object when he can produce its *technical description*. The *mental image* of the object remains in the mental sphere of the cognitive subject until the latter can communicate it through behavior.

   Repeated observations of *objects*, which are known through their [des O], along with their comparison using the observational dimensions selected in the Obs-UN trigger the process of *analogical generalization* that leads to *understanding*, since the so-called idea of the object is produced which does not correspond to particular *mental objects* but to a conceptual *abstraction* ("the square," "the rectangle," "the triangle," etc.) deriving from *analogical retention* of what is considered *common*, useful to *incorporate,* and *unifying* regarding the various mental descriptions

MENTAL OPERATION	FORMALIZATION	SYMBOLISM
Image of O	Distinction in the Obs-UN	O
Configuration of O	Technical description	[des O]
Idea, concept, abstraction of O*	Technical definition	[def O*]
Understanding	Meaningful technical definition	[min def O*]

Fig. 9.9 Mental representations of objects of observation

(number of sides, size of angles, length of sides, etc.). The "idea" of the *general object* (also referred to as *concept, notion, generalization, abstraction, intension*, etc.) corresponds to the object O* of the *analogical generalization* specified by [des O*].

As we have just observed, *simplifying* mental activity results in neglecting in the *technical description* of objects and the *technical definition* of concepts many dimensions with "little weight for the *retention, conservation,* and *evocation* of the "idea" of the mental objects. This simplification leads to the formation of "essential ideas" formed and conserved using a minimum number of relevant dimensions. We can refer to this "essential idea" as the "understanding" of the object, and in the mental sphere in our formal system this corresponds to the *meaningful technical definition:* [min def O*].

Thus we understand "square" by generally considering only the descriptive values for the sides (four, equal in length) and angles (all equal), and not by observing the absolute dimension or the color of the sides and surface area. The translation of the symbolic structure proposed for the terms *image, configuration, idea,* and *understanding* is summarized in Fig. 9.9.

As our experience shows, distinct objects never appear to the mind. The perception of a distinct object of observation is nothing other than the result of an activity of selection-attention undertaken by the observer in order to isolate that object in an expressively broader context (recall the distinction between united and separated objects in Sect. 9.5). The broad context in which the object of observation is identified can be referred to as "landscape." A *landscape* is not a composite object since its components are distinctly observed even though connected in terms of belonging to that scenario. Nevertheless, a *landscape* (or scenario) can be considered a composite object whenever the observation involves the entire *landscape* as a unit.

If M objects: $O^1, O^2, \ldots, O^M$, constitute a *landscape* "L," then we can write:

$$[L] = \left\{ \left[\text{des } O^1 \right] \cup \left[\text{des } O^2 \right] \cup \ldots \left[\text{des } O^M \right] \right\} \tag{9.7}$$

If, for all the objects composing a scenario [L], we emphasize the time dimension "duration" (as perceived by humans), defined in meaningful terms, then the entire scenario can also be considered to be observed in the time dimension. This scenario *evolves* through changes in the number and type of objects it is made up of

at any moment. I shall refer to "environment," [E], as a scenario in which the duration of observation, T, is also defined. An environment is *historical* if the final moment in T corresponds to the actual moment of the observation or to the future moment in cases where the date of the observation coincides with the start of T. For the sake of convention, an actual or present environment is when the moment of observation is included in T.

Therefore, to an observer every object can be part of a landscape and be observed in an historical, present, or future environment.

9.10   The Signification Process: Signs, Denomination, Signifier, Seme, and Sememe

In order to analyze the process by which the *mind* forms knowledge—that is, the *thought process*—it was necessary to express myself through a *language*. However, it should be clear that until now, though having to express myself linguistically, I have described *basic mental operations* deriving from the mind's ability to calculate, compare, and associate differences, assuming the mind has not yet developed a *language* to share its knowledge with other minds; that is, to *communicate* its *thought content* with other cognitive subjects. The *mental knowledge* developed by a cognitive subject is, so to speak, a "private" phenomenon which can be perceived from "outside the mind" only if that subject is able to produce some form of *behavior*, in particular a *meaningful behavior*, to communicate the knowledge it has acquired.

The perception of danger and fear in a dog can be deduced by observing its "tail between its legs"; the pleasure felt by a male peacock upon seeing a female one can be deduced by observing how it displays its tail; the lion's desire to attack a prey can be deduced by its attack position, while the horns pointed toward the lion reveal the mental activity of the buffalo, which is turned to its defense on perceiving the predator. All "higher-order" animals, in particular quadrupeds, cetaceans, fish, birds, etc., even crustaceans and mollusks, have developed some form of signaling behavior to express to other individuals some thought content. I am certain the reader will agree that the "whale's song," that is, the catalog of sounds produced by whales, dolphins, and porpoises, is nothing other than behavior from the emission of sound waves through appropriate effectors in order to communicate to other members of the species. Humans have developed a very powerful communicative form of behavior through the formation and use of numerous *languages* which makes it possible to efficiently communicate the thought content the mind has developed (Sect. 9.11).

The analysis of the structure and function of languages goes beyond the scope of this book. In order to complete this brief "operational translation" of Bateson's model, it is nevertheless useful to at least examine the *denomination* process through which the "mind" associates to its *mental content* (that is, the

objects of observation, descriptions and *definitions*) other *conventional objects of communication*: the intentional "signs" produced through its own behavior (speaking, gesturing, writing, pressing keys, etc.), using these "signs" to then communicate to other minds its own "private" *mental objects* by transmitting the signs in the form of *signals* emitted by some specialized effector.

In short, the *process of communicating* thought content, T_A, from a cognitive subject A (transmitting subject) to subject B (recipient) assumes that:

1. A associates to T_A the sign ST_A it produces; that is, it *denominates* T_A by the sign ST_A.
2. A transforms ST_A into a signal, σT_A, obtained from one of its effectors (indicating with the hand, transmitting sound, etc.).
3. the signal σT_A reaches B.
4. B, through σT_A, recognizes the sign ST_A that A intended to transmit.
5. B, through ST_A, forms in its own mind the thought content *denominated* by ST_B.
6. $ST_B = ST_A$.

Condition (6) states that the communication is successful when B "recreates in its mind" the same thought content that A intended to transmit to it. In the simplest cases the success of the communication can be perceived both by A (transmitter), evaluating the behavior of B (appropriate response by B to A's question; appropriate execution by B of A's order, etc.), and by B (recipient), evaluating the correct effects from its own behavior, which is carried out by interpreting ST_A.

This interpretation of the communication process, as an act involving the transmission of thought content by the transmitter (in particular *technical descriptions* and definitions), through "signs" that allow the mind of the receiver to reconstruct the former's thought content, is perfectly in line with the autopoietic perspective of cognitive systems that produce linguistic interactions.

> Due to the nature of the cognitive process and the function of the linguistic interactions, we cannot say anything about that which is independent of us and with which we cannot interact; to do that would imply a description and a description as a mode of conduct represents only relations given in interactions. [. . .] it follows that reality as a universe of independent entities about which we can talk is, necessarily, a fiction of the purely descriptive domain, and that we should in fact apply the notion of reality to this very domain of descriptions in which we, the describing system, interact with our descriptions as if with independent entities. This change in the notion of reality must be properly understood. We are used to talking about reality orienting each other through linguistic interactions to what we deem are sensory experiences of concrete entities, but which have turned out to be, as are thoughts and descriptions, states of relative activity between neurons that generate new descriptions [. . .] we recognize that we, as thinking systems, live in a domain of descriptions, as has already been indicated by Berkeley, and that through descriptions we can indefinitely increase the complexity of our cognitive domain (Maturana and Varela 1980, pp. 52–53).

It is evident that the *denomination* process is fundamental for all communication. Recalling Bateson's motto: "The map is not the territory, and the name is not the thing named," and also Saussure, for whom "The linguistic sign unites, not a thing and a name, but a concept and a sound-image" (Saussure 1916, p. 66), I shall

define *denomination* as the process by which a *sign* (a conventionally accepted name, in particular) is assigned to a concept.

> Some people regard language, when reduced to its elements, as a naming-process only—a list of words, each corresponding to the thing that it names. ... This conception is open to criticism at several points. It assumes that ready-made ideas exist before words; it does not tell us whether a name is vocal or psychological in nature ...; finally, it lets us assume that the linking of a name and a thing is a very simple operation—an assumption that is anything but true. But this is a rather naive approach [...] The linguistic sign unites, not a thing and a name, but a concept and a sound-image. The latter is not the material sound, a purely physical thing, but the psychological imprint of the sound (Saussure 1916, pp. 65, 66).

> The bond between the signifier and the signified is arbitrary. Since I mean by sign the whole that results from the associating of the signifier with the signified, I can simply say: *the linguistic sign is arbitrary* (ibidem, p. 67).

Only when the baby's grandmother in the example in Sect. 9.7 pronounces the vocal sign "pear" by pointing at several *pears* does she (ostensibly) *denominate* using this vocal sign "pear," a fruit the baby has mentally observed; that is, which he has described several times, thereby forming a concept. When the baby wants a pear (*definition*) as fruit he will use the sign "pear" he has heard, *trying to use the correct pronunciation*. Observing his mother's smile, he will note her "sign" of joy in hearing her son denominate that type of fruit, and the child will thus be convinced that using the vocal sign "pear" is very effective for transmitting to others his thought content (*definition*) regarding the nature of the desired fruit.

Signs can be used in three types of *denomination*:

(a) *Proper denomination* matches a *sign* to a *technical description* of a single object "O," and that sign becomes the *proper name* of the *described* object, the only one which can be denoted by that description and which represents the *signified* of S.

$$\textit{Proper denomination} \text{ of } [\text{S denoting "O"}] = [\text{des O}] \qquad (9.8)$$

(b) *Intensive common denomination* matches a sign to a *technical definition*. That sign becomes the *common name* for all those objects, denotable by that definition, which constitute the *signified* of S.

$$\textit{Intensive common denomination} \text{ of } [\text{S denoting O}^*] = [\text{def O}^*] \qquad (9.9)$$

(c) *Extensive common denomination*, which explicitly indicates all objects that can be denominated by S.

$$\textit{Extensive common denomination of} [\text{S denoting O}^*]$$
$$= [\text{des O}_1, \ldots \text{des O}_N] \qquad (9.10)$$

The objects which a sign S can denominate and indicate—individual objects or concepts—form the "meaning of S."

$$\text{\textit{Meaning of}}\,[\text{S denominating O, or O}^*] = [\text{des O}]\,\text{or}\,[\text{def O}^*] \qquad (9.11)$$

The *meaning* is conventional for a group (or social context) and refers not to the *objects indicated* but to the *indicator signs*. Imagine the baby's surprise when, during a trip to Turkey, he asks for a pear using the vocal sign "pear," but without being understood.

When I pronounce the vocal sign "university" it is totally different from the one pronounced by my daughter, my grandchild, my neighbor, or any of the readers. We all use different inflections, tones, and we all have different pitches, length of pronunciation, expressions, being able to pronounce the vocal sign either by smiling or crying, with clenched teeth or an open mouth, with surprise or anger, etc. And yet our mind, when considering the different signs for "university," as *objects* that have been perceived several times, compares them and quickly finds the similarity by creating the *technical definition* of the objects-signs-vowels: "university." The same process occurs reading the graphical signs (the underlined words are originally colored, while they appear to the reader as black and white):

University, UNIVERSITY, University, **university**, *University*, university, University, UNIVERSITY etc.

These are all different *signs*, but our mind (due to our basic school instruction) has soon become able to form a *generalizing analogical abstraction* in which, and despite the evident differences in color and font, all the graphical signs contain the same letters and are associated with the same content we are all familiar with.

The set of all "signs" having the same *signified* mental object (description or definition) is the *signifier of the sign* and corresponds to the *technical definition* of the sign:

$$\text{\textit{Signifier of the sign}}\,[\text{S}^*] = [\text{def S}^*]. \qquad (9.12)$$

As a result, every sign which is part of a *signifier* is appropriate for indicating any of the components of the corresponding *signified*; likewise, any element of a *signified* can be indicated by any sign of a corresponding *signifier*. The correspondence (denoted by $\leftrightarrow$) of the *signifier* (9.12) to what is *signified*, (9.8), (9.9), or (9.10), represents a *seme*, in Prieto's sense of the term (1966), and the convention adopted by a collectivity to achieve this correspondence is the *semic code*.

$$\text{\textit{Seme}} = [\text{def S}^*] \leftrightarrow [\text{def O}^*],\,\text{by means of a}\,\text{\textit{semic code}}. \qquad (9.13)$$

The *sememe* of a sign is defined as the set of all *interpretants* of the *meaning* of the sign in a specified language (Prieto 1966).

$$\text{\textit{Sememe of}}\,[\text{def S}^*] = \text{Interpretants of}\,\text{\textit{Meaning of}}\,[\text{S denominating O, or O}^*]$$
$$(9.14)$$
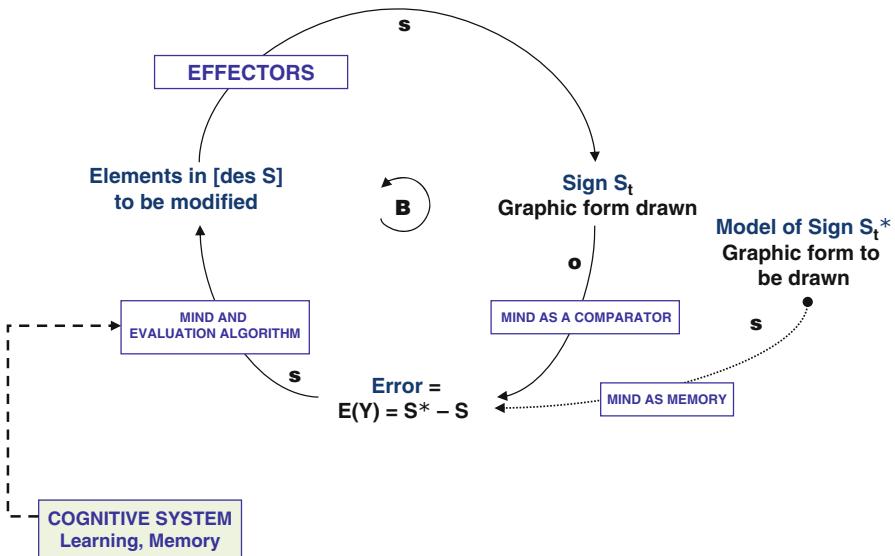
Fig. 9.10 *Ring* for controlling the correct form of a sign

In order to better specify the three concepts, signifier, seme, and sememe, note that the graphical sign "dog" denotes a *concept* (technical definition) which can indicate different specific animals that the mind recognizes as dogs:

["Fido," "Buck," "Smurf," "Ball," "Black," "Samson," etc.] = Extensive common denomination = *extensive meaning* of the sign "dog."

The same meaning (the set of the observed dogs) can also be evoked by the signs:

[dog, *dog, dog,* DOG, *dog*, DOG, **dog**, DOG, dOg, DoG, d o g , D O G, **D ○ G**, etc.] = *Signifier* of the sign "dog".

["our most faithful friend," "the faithful guardian of our home," "the most intelligent of pets," in addition to other expressions that designate "dog"] = *sememe* of [def S*].

It is clear that the introduction of *signs* and *semi* require the use of multiple Recognizing and Identifying Control Systems (Sect. 3.2) that operate at very high speed and over a great many cycles. Figure 9.10 presents the *Ring* for "constructing" the correct graphical form of a *written sign*, S, or the correct pronunciation, if S is a vocal sign, assuming that the mind has already formed the *Signifier of* [S*], as defined in (9.12).

When the cognitive system that arises from applying the Ring in Fig. 9.10 succeeds in forming and communicating the *sign*, learning to do so more and
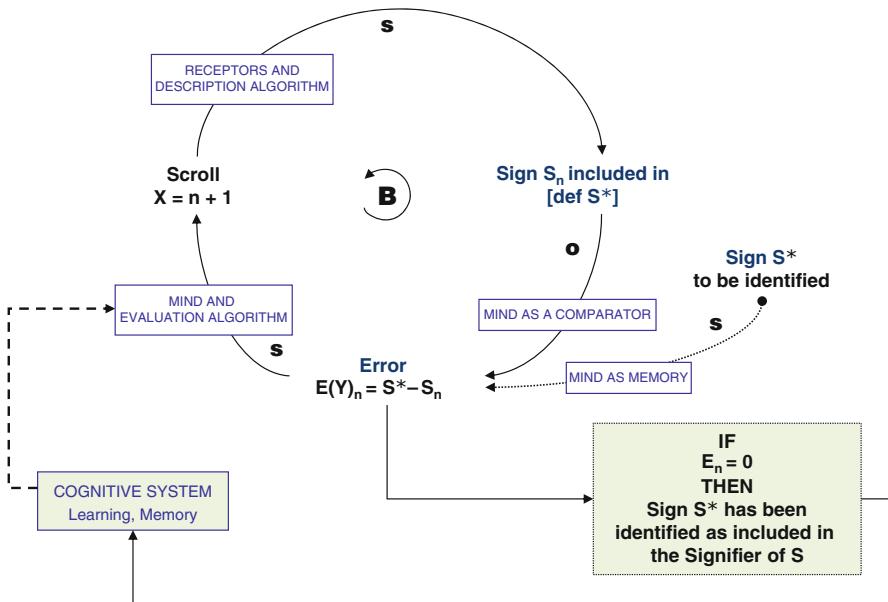
Fig. 9.11 The *Ring* to identify a sign belonging to a signifier

more quickly, the receiver must be able to *identify* and *interpret* it. The identification of the sign represents the mental operation which identifies the *Signifier of the sign* = [def S*], to whom the sign belongs, as illustrated in (9.12). The Identifying Control Systems are thus activated, which are similar to those in the model in Fig. 9.11.

Once the sign is identified the recipient must *interpret* it; that is, determine its meaning. From (9.8) we know that, in order to determine the significance of the sign S representing a proper name, each recipient must recognize or identify the object determined by the [des O] or [min des O]; if the sign S indicates a concept, such as in (9.9), then every recipient must identify or determine the correct [def O*] or [min def O*], as defined in Sect. 9.8.

Having learned from earliest infancy to manage the Recognizing and Identifying *Rings*, it is not easy for us to realize how intensively and quickly the *Rings* operate that choose the correct meaning from our Obs-UN. One such *Ring* is illustrated in the model in Fig. 9.12.

The above considerations can be easily understood if we turn again to the example of the baby who knows the fruit and vocal signs that designate the concepts (technical definitions) of the various fruit, recognizing their similarity even though they are described differently. Thus, consider the case regarding the recognition of the signs of the following propositions (strings of signs):

これらは兆候です , 이러한 징후입니다
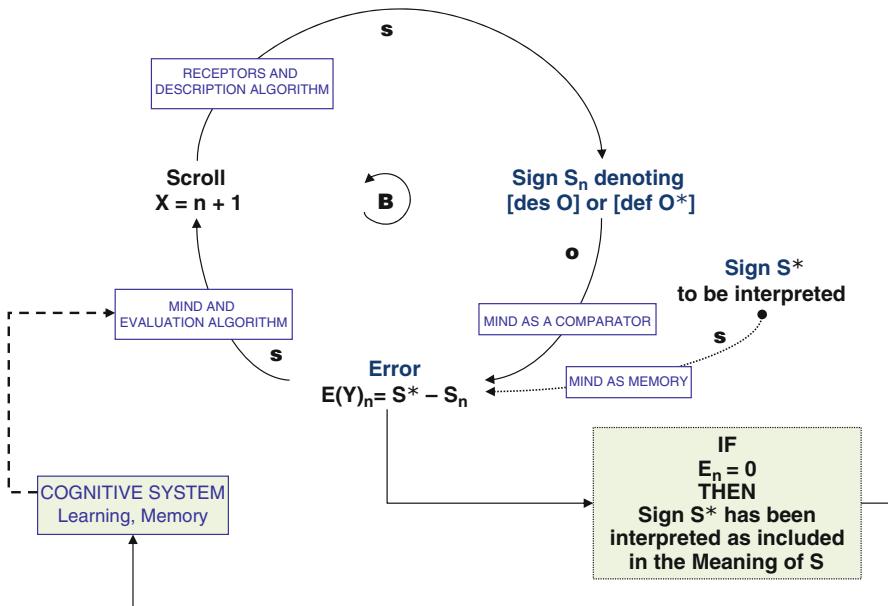
Fig. 9.12 *Ring* for interpreting the meaning of an identified sign

I am not able to identify the signs, and thus I cannot interpret them, even if I am sure the mind of a Japanese or Korean person would immediately recognize the *signifier* of the signs and interpret them by forming the sentence: "these are signs."

9.11   Languages and Signs

Following Prieto, *semiology* is the science that studies human behavior regarding the attribution of *semes*, which represents a broader meaning than the original one attributed by Saussure. The field of *semiology* intersects with both the *linguistics* field, which studies linguistic signs and languages, and *semiotics*, the discipline that studies linguistic and nonlinguistic "signs" (Nöth 1995).

> […] without doubt verbal language is the most powerful semiotic device that man has invented; […] nevertheless other devices exist, covering portions of a general semantic space that verbal language does not. So that even though this latter is the more powerful, it does not totally satisfy the effability requirement; in order to be so powerful it must often be helped along by other semiotic systems which add to its power. One can hardly conceive of a world in which certain beings communicate without verbal language, restricting themselves to gestures, objects, unshaped sounds, tunes, or tap dancing; but it is equally hard to conceive of a world in which certain beings use only words; when considering … the labor of mentioning states of the world, i.e. of referring signs to things (in which words are so intertwined with gestural pointers and objects taken as ostensive signs), one quickly realizes that in a world ruled only by words it would be impossible to mention things (Eco 1976, p. 174).

I propose to define *language* as a system of *intentional semes* and *sememes*, codified by a collectivity, through which men can attempt *linguistic communication* involving any thought content whatsoever, by producing a system of *semic actions*. There are various types of languages, which differ according to the signs they use and the "codes" through which the *denomination* process unfolds. Some are specialized languages, employing, for example, logical, mathematical, graphical, or iconic signs. The most extensive language is the *natural verbal* one, through which the human mind can communicate any type of thought to another mind.

> [This can be expressed] . . . by saying that every theory of signification and communication has only one primary object, i.e. verbal language, all other languages being imperfect approximations to its capacities and therefore constituting peripheral and impure instances of semiotic devices. Thus verbal language could be defined as the *primary modeling system,* the others being only "secondary," derivative (and partial) translations of some of its devices . . . Or it could be defined as the primary way in which man specularly translates his thoughts, speaking and thinking being a privileged area of a semiotic enquiry, so that linguistics is not only the most important branch of semiotics but also the model for every semiotic activity; semiotics as a whole thus becomes no more than a derivation from linguistics (Eco 1976, p. 172).

Every language can be viewed in instrumental terms. A fundamental function of languages, no matter what their form (written, spoken, gestural, mimicking, ritual, iconic, and so on) is to enable communication to occur between individuals in order for them to share thought content.

In his view of social systems as autopoietic systems, Niklas Luhman assigns language a fundamental role, since it permits communication among individuals and the generation of orders and progress (Luhman 1995). In his book, Theory of Society (1997), he pays particular attention to language, claiming that the words (of the spoken language) have no precise significance if considered alone, but rather serve the function of producing sentences with meaning and, above all, increasingly newer sentences.

> Speech possesses its very own *form*. As a form with two sides it emerges from the difference between *sound* and *meaning* . . . Spoken communication is the processing of meaning within the medium of sound. . . . In order to construct the difference between medium and form within speech itself, the medial substrate of speech, the difference between sound and meaning, must be *underspecified*. Without under specification there would be nothing left to say because everything would already have been said. This problem is solved with the difference between words and sentences. Words are also constellations of sounds with meaning; but they do not determine the sentences that they can produce when combined. . . . With the help of speech, one may say *something that has never been said before* (italics in original) (Luhmann 1997, pp. 213–215; translation by Daniel Lee 2000, p. 326).

Thought contains basic content—[des O] and [def O*]—and composite content, which derive from a union of basic content in time and space. Thought is continuous over time and discontinuous regarding its content. It is continuous in that during wakefulness, it flows uninterruptedly; it is discontinuous in that thought content can suddenly change, shifting from one mix of objects and concepts to a new one when called for by communication and the survival of the cognitive

subject. Language not only tries to express thought content by using appropriate *semes* that adapt to the flow and necessities of the cognitive subject, but it also tries to express such content more precisely. Therefore, language is revealed as a *flow of signs* denoting successions of *thought content*, even in a memorized form. Despite this variety, the thought content language can transmit can be reduced to the following categories:

1. *Expressions of judgments* and *emotions*, understood also as expressions of opinions and evaluations, impressions, indications of kindness, beauty, amazement, annoyance, happiness or unhappiness, and so on ("What a beautiful rose!").
2. *Declarations of intentions* to undertake certain actions or types of behavior, positive and negative ("I desire to pick the rose before it withers").
3. *Orders*, understood as manifesting a desire, or a necessity, to a subject that he behave in a certain way ("You go pick the rose, but be careful not to leave the stem too short").
4. *Questioning*, expressed as requests for thought content and answers ("Should I also pick the tulips?").
5. *Answers to questions* and *orders*; that is, the revelation of thought content that has been requested ("Ok. I'll go then and pick the rose"; "No, leave the tulips where they are").
6. *Descriptions* and *procedures*; that is, the results from observing objects and portions of reality and the procedures for obtaining certain objects as the result of the application of the rules established by the procedure.
7. *Information*, or *specific data* useful in carrying out operations or activities.
8. *Argumentation*, through which the cognitive unit tries to judge the truth of certain statements (answers, descriptions or information), and seeks the explanation of the phenomena observed (Sect. 9.15.2).

Functions 1–5 mainly communicate the "internal states" of the cognitive subject; functions 6 and 7 communicate the "states of the world," that is, the states of the variables that describe or define objects assumed observable by both parties involved in the communication. Function 8 tries to arrive at the *conviction* that there exists a correspondence between the significance inferable from a statement and the "states of the world" the statement indicates; in particular, it tries to derive *explanations* of the "states of the world," deriving these from other "states of the world" which are observed or assumed or from a theory on the states of the world.

The distinction among the eight functions is in part subjective, since even the evaluations, questions, and orders can be interpreted as manifestations of information regarding the desires or sensations of the communicating subject; at the very least they have an informational content. The argumentations, on the other hand, make use of judgments, above all basing themselves on a system of information; they can serve as the premise for an order. Example: information: "Our friends expect us for dinner in 20 minutes"; expression: "Aren't you ready yet?!?"; information: "It takes 20 minutes to get there by car"; judgment: "Come on! You're always late because you're too slow"; question: "You want to get me in

trouble?"; order: "Hurry up then!"; argumentation: "Otherwise we'll arrive too late"; answer: "I'm not the one whose slow; it's you who are stressing me out!"; argumentation: "And it's not true it will take us 20 minutes, because we'll take the ring road!."

Having distinguished among these basic types of thought content leads us to conclude that the fundamental function of languages, that of communicating thought content, can be divided into *specific functions*: the *expressive, declarative, imperative, inquisitive, replicative, descriptive, informative, and argumentative* functions.

The classification proposed conforms to the language functions analyzed in terms of transmissible thought content, and can be traced to the famous linguistics classification proposed by the scholar in *phonology*, Roman Jakobson (1970). Jakobson analyzes the functional aspects of linguistic acts in the *context of the communication process*, distinguishing among[1]:

1. The "expressive" or "emotive" function: "This tends to give the impression of a certain emotion, true or simulated."
2. The "conative" function: "The conative function is expressed in its purest grammatical form in the vocative and imperative case which, from a syntactical and morphological, and even often phonological point of view, diverges from the nominal and verbal categories."
3. The "phatic" function: "There are messages that serve essentially to establish, extend, prolong, or interrupt communication, to check if the circuits are working, attract the attention of the interlocutor or to ensure it does not fade."
4. The "metalinguistic" function: "Whenever the sender and/or recipient deem it necessary to check whether or not they are properly using the same code, the discourse centers on the code: it fulfills a metalinguistic function."
5. The "poetic" function: "The poetic function is not the sole function in the art of language, merely the dominant, decisive one, despite the fact that in other verbal activities it only plays a subsidiary, accessory role."
6. The "referential" function: informative propositions that communicate contextual facts, or "states of the world."

---

[1] Below is the description of the functions in their original version in French, in Roman Jakobson (1970), pp. 213–222: "Fonction dite "expressive ou émotive" ("Elle tend à donner l'impression d'une certame émotion, vraie ou feinte"); "conative" ("la fonction conative, trouve son expression grammaticale la plus pure dans le vocatif et l'impératif, qui, du point de vue syntaxique, morphologique, et souvent méme phonologique, s'écartent des autres catégories nominales et verbales"); "phatique" ("Il y a des messages qui servent essentiellement établir, prolonger ou interrompre la communication, à vérifier si le circuit fonctionne, attirer l'attention de l'interlocuteur ou à s'assurer qu'elle ne se relàche pas"); "metalinguistique" ("Chaque fois que le destinateur et/ou le destinataire jugent nécessaire de vérifier s'ils utilisent bien le même code, le discours est centre sur le code: il remplit une fonction métalinguistique"); poetica ("La fonction poétique n'est pas la seule fonction de l'art du langage, elle en est seulement la fonction dominante, determinante, cependant que dans les autres activités verbales elle ne joue qu'un róle subsidiaire, accessoire"), "referentielle"."

It is not difficult to relate these six functions to the eight types of thought content that can be communicated.

Language can also be used *pragmatically* (phatic function) even simply for *ceremonial aims* ("Good morning," "Welcome," "Pleasure to see you," etc.), and can carry out a *self-referencing* function by indicating to the receiving subject that the sender wants to start, continue, or stop communicating with him ("Hey!," "What rainy weather," "Over and out"). In the case of the self-referencing function we can distinguish the *indexical* function, where the language, especially oral, serves to express the subject's personality (use of poetic or academic phrases), emotional state (tone of voice, stuttering, incisive or confused writing, or containing a number of corrections, etc.), or social class (types of signs used, the speed with which the signs are produced, the use of slang expressions, etc.). Finally, language is also an instrument for lessening our anxiety ("Come on!!! How much time does it take the bus to get here!").

Usually communication serves to transmit relatively complex thought content and to allow interactions among the various subjects or groups of subjects. Apart from simple cases ("Stop," "Be careful," "Enough!," etc.) several *semic* actions are needed to transmit systems of signs variously arranged, even when they are chosen by universes of different signs (for example, the oral word can be accompanied by signs in the miming-gestural language).

The semiological and linguistic study of the way systems of signs are arranged and form in languages can be undertaken following three approaches:

1. The *semantic* study, which specifically studies the relation between signs and meanings for both the senders and recipients.
2. The *syntactic* study, which analyzes the relations between signs and the possibility of coherently combining them.
3. The *pragmatic* study, which considers the ways in which the linguistic signs are used in practice within a given "community" of individuals characterized by a given culture in order to "do something" to achieve results or objectives. For example, the sign "stop" written in words or on a sign at the end of the street, the sign "Come at 9 a.m.," or the sign "I understand. I'm coming" clearly are sent to produce or declare the effects of behavior and they produce or declare the effects of behavior if the meaning of the sign is recognized by those receiving it. This ability of signs (and thus of languages) is known as *performativity*, a term introduced by John Langshaw Austin, a language philosopher who has studied the pragmatic effects of signs and languages.

> It is worthy of note that, as I am told, in the American law of evidence a report of what someone else said is admitted as evidence if what he said is an utterance of our performative kind: because this is regarded as a report, not so much of something he said, as which it would he hear-say and not admissible as evidence, but rather as something he did, an action of his. This coincides very well with our initial feelings about performatives (Austin 1962, in Bial 2004, p. 145)

It is evident that signs are evaluated in the pragmatic sphere in a community that shares the *semic code*. In particular, *performativity* characterizes juridical language, which by definition is normative.

I would like to conclude this brief synthesis by considering the types of signs that must characterize all languages in order for these to carry out their communicative function, thus allowing the thought content to be *enunciated* through a succession of signs forming a *proposition*. A *proposition* is a sequence of elementary signs with a complete meaning, in that the signs serve to enunciate, express, and transmit the content of a defined thought (expressions, orders, questions, information, argumentation, etc.). Let us assume that those signs are determined by a unique *semic code* that is conventionally accepted by both the speaker and listener. In order to express complex thought content we use a *statement* made up of a sequence of propositions that express a system of meanings representing the overall meaning. The meaning of a proposition or a statement represents the *message* the sender wishes to transmit to the recipient, which falls under one of the eight functions examined above.

Without entering into a typological distinction, which can be found in any dictionary and grammar book in any language, the signs (in particular, the vocables, words, literary terms) that make up the propositions and statements can be divided into two *main* classes and two *accessory* ones, according to their main function.

Above all, there are the signs of a given *signifier*, whose main function is to denote objects (nouns) or actions (verbs) and their characteristics (attributes) according to how the observer forms these in his mind. For simplicity's sake, considering only the signs-words in the written language, note that the sign "Piero" is a proper name that indicates the *idea* of a generic individual with this name. In order to indicate a specific "Piero" we must use signs that designate some elements in its *technical description*, in particular in the *meaningful technical description*. If in my mind I am thinking of "me" and wish to communicate who I am, I must use words that indicate dimensional states that distinguish my [des PIERO]. For example, "Mella," to indicate the state of the "surname" dimension; "male" to indicate the state of the dimension "sex"; "born in Italy," to indicate the state of the dimension "place of birth"; "tall ...," "hair color ...," etc. For a *meaningful* technical description it could be enough to produce the phrase "Piero Mella, professor at Pavia University, in Italy." If my son were to tell a friend that he was referring to [des PIERO] he would simply say: "my father."

Though we seldom realize it, most of the words we use indicate *dimensions* (height, speed, weight, color, license plate, etc.) or attributes and dimensional states (180 cm, 60 km/h, 80 k, green, JK 335 OP, etc.). Even *actions* expressed by signs-verbs are part of the descriptions or definitions: "Piero took a taxi," "Mary runs 2 miles every day," "I ate a steak with fries," etc., which describe observed objects. We can also include in this category the signs that indicate an intention or judgment: "I'm going home," "I don't think I'll take a taxi," "It's been a terrible day and I'm tired." Even questions and orders must be expressed with signs that indicate objects of thought referring to descriptions or definitions: "Go get the pen on my table," I can't find it," "Are you blind? It's right in front of your eyes!."

In order to transmit thought content through propositions and statements, composed of the signs from the preceding category, a *second* category of signs is necessary which functions in order to *give orders to the mind*, in particular to its

attention capacity, so as to evoke the "mental operations" needed to arrange the *objects of thought* in an appropriate "order" in the context of the linguistic state-ments. I am referring to the *ordering* or *correlating* signs for the meanings of other signs, or, in general, to the signs that evoke the *mental operations* that give "orders" to the meanings of other signs from the same class.

In written English, for example, the sign "...on..." does not *denominate* any object of observation but merely *evokes* the *mental operation* "overlaying" needed to order and correlate the *objects of thought* indicated by "....." For example, the proposition "the bottle is *on* the floor" has a complete meaning, since it allows us to state an acceptable mental object; "the floor is *on* the bottle," on the other hand, forces us to evoke a concept whose meaning is unacceptable. There are many and varied linguistic signs that designate *mental orders* in order to allow the mind to form the objects of though indicated by the language. I shall leave the reader to recall the numerous classes of signs of this type; here I shall mention only the most obvious ones:

1. The numerous types of signs of *conjunction*: "...and...," "...or...," "...either...," "...neither...," "and...and...," "either...or...," "neither... nor...," "both...and...," "...also...," etc.
2. The *articles,* "the..." and "a..."; in many languages these are specified by grammatical category: male or female, singular or plural, relative to the words that express mental objects.
3. The *prepositions*, "of...," "to...," "from...," "in...," "with...," "on ...," "between...," etc.
4. All the various types of *adverbs*, including *adverbs of time* ("still," "later," "yesterday," etc.); *adverbs of quantity* ("many," "little," "too much," etc.); *adverbs of quality* ("slowly," "carefully," "suddenly," etc.); *adverbs of place* ("above," "below," "in front of," "behind," "right," "left," etc.).
5. The signs to indicate singular or plural, male or female; for example, in English "words" order the mind to identify a group of words; the statement "my cars is red" and "my child are at school" stop someone receiving them (outside of any particular context) from identifying any meaning, since the plural and singular expressions contrast with the meaning of the verb.

In order to see how the signs of the preceding categories send *orders* to our mind, consider what is evoked by the following statements: "Carla grabbed the pencil and placed it above...," "Carla grabbed the pencil and placed it above the ...," or "Carla grabbed the pencil and placed it above our ...." I could continue on with this example, but a moment of thought will surely enable the reader to see that every sentence spurs the mind to complete the image the speaker is creating. The role of the signs sending orders to the mind and "creating" the meaning of a statement was pointed out by Silvio Ceccato, an Italian linguist and cybernetician who, in his book entitled *Course in Operational Linguistics* (1969) demonstrated that the *ordering* signs are finite in number and, through appropriate *sequences of attentional acts*, their operational meaning in the construction and interpretation of linguistic expres-sions can be defined. Ceccato proposes a simple experiment: place the tips of the

fingers of one hand on the edge of a desk. Reflect on what you "feel" if I tell you: "you are touching *the front part* of the desk," or alternatively: "you are touching *the back part* of the desk." Your attention is diversely affected by each sentence and, in fact, you think you are touching what is "ordered" by the signs "front part" and "back part" with a shift in attention.

Languages usually include a third class of signs: the so-called signs of *punctuation* (".", ",", ":", "/", etc.), whose function is to separate or order the linguistic signs arranged in sequence in order to allow the mind receiving the message to correctly identify the succession of meanings of the words that make up the statement or proposition. A fourth class includes the *functional signs* (for example, "?", "!", "!?", etc.), whose function is to evoke the linguistic subfunction chosen by the sender in transmitting a statement. The signs of the last two classes are considered *accessory* since their use can be avoided by replacing them with the signs from the main classes; for example, "Come here!" can be interpreted in the written language as "I order you to come to me." A proposition (and thus a statement) will be composed of signs from the four classes indicated above, arranged according to the conventional order set by the code that attributes meaning to the terms used.

9.12   Factual Truth and Falseness

Among the eight functions of languages, of particular importance is the *argumentative* one, since it carries out linguistic transmissions on the "state of the world" a mind observes and desires to transmit to another mind through language. Following Copi and Cohen (2011), we shall define a *declarative proposition*, or *statement*, in any language as a sequence of basic signs capable of expressing a thought content that can be ascertained to be "true" or "false" using some *conventional procedure*.

> We know that every statement is either true or false. Therefore we say that every statement has a truth value, where the *truth value* of a true statement is *true*, and the truth value of a false statement is *false*. Using this concept, we can divide compound statements into two distinct categories, according to whether the truth value of the compound statement is determined wholly by the truth values of its components, or is determined by anything other than the truth values of its components (Copi and Cohen 2011, p. 290).

Verification and *falsification* are the cognitive operations that *verify* whether or not a *declarative statement* regarding differences, descriptions, definitions, or cognitive procedure (Sect. 9.11) has an observable or derivable meaning which is in accordance with the observations on the "states of the world." Thus, the truth does not refer to the "states of the world" but to the *statements* on the states of the world. As *truth* derives from a *mental process* of verification or falsification, it is always relative, at least as regards the instruments for observing the state of the world and the language used to construct the statement.

I wish to hypothesize about how the "Batesonian mind" can ascertain the truth. Let us suppose that a declarative proposition, *E*, transmitted from Alfa to Beta in a

given language, asserts that the difference: $d_n(A^*) \in$ [def A*] is *true;* for example, "Snow is white" states that in [def A*] = [def "snow*"], $d_n(A^*)$ = "color of snow = white." The proposition "Snow is white," composed and transmitted by Alpha, is *true* for Beta if the latter can, using his own procedure (which is not necessarily similar to that used by Alfa), construct [def A*], thereby determining $d_n(A^*)$ and verifying it belongs to the definition; otherwise the proposition is *false*.

> But one might object that, ultimately, truth is a matter of using and accepting a sentence as an adequate description of a state of affairs. Thus, "Snow is white" is true if and only if we are prepared to use and accept that sentence to describe a property that snow in fact has. "Snow is black" is a misassignment; it is false, because we are not prepared to use and accept that sentence as a description of snow [*more correctly: a description of a dimensional state of any object we denote as "snow"* (author's note)]. But with "Snow is marble" we may begin to hesitate; perhaps in certain circumstances, it *is* a true metaphorical description (Ankersmit and Mooij 1993, p. 78).

Adopting Bateson's concept of mind, let us consider Alfred Tarski's rule (1944) for truth: "'Snow is white', is true if and only if snow is white" (Tarski 1944, p. 342). The formal statement "Snow is white" is *true* for Beta if Beta *observes* that, in objects where [des snow] $\subseteq$ [def snow*], the dimension D_{COLOR} always assumes the state $[d_{COLOR}(snow) = \text{" white "}] \in$ [def white*]. According to the procedure examined in Sect. 9.7, *truth* thus depends on the processes of *distinction* (white?) *understanding* (snow? color?) and *recognition* ("this" is snow and its color is white?) and assumes a reliable *process of determination* (What does "snow" mean? What does "white" mean in a chromatic scale?). These observations are perfectly in line with Tarski's general definition of truth: "A sentence is true if it is satisfied by all objects and false otherwise" (Tarski 1944, p. 353).

Alfa's statement to Beta that $[\Delta d_n(O^*)] \subset$ [def O*] (for example, "men are mortal") or that [des O_m] $\subset$ [def O*] (for example, "Piero is a man") is *true* if Beta is able to undertake a cognitive procedure that can construct [def O*], that is, "man," and can recognize that $[\Delta d_n(O *)]$, that is, "mortal," belongs to it and identify [des O_m] = Piero as an element of [def O*]. In conclusion, the statement "Piero is a man" is *true* if [des PIERO] $\subset$ [def MAN].

The statement "The fourth root of 100,000 is 17,7827941" is true if the [def SQUARE ROOT] includes a calculation procedure that, when applied, provides the declared value.

The concept of truth and falsehood, made operational by using Bateson's model, is *factual* in nature because the truth is conceived of as a *correspondence* between the meaning of the proposition and what the mind can observe. However, the *argumentative function* of language also tries to determine the truth or falsehood of *composite statements* through the relations of truth and falsehood in the specific, or "atomic" statements, taking account of the meaning of the connective signs that link them together. I cannot deal here with the topic of the search for formal truth and falsehood, since this topic is typical of the logic, in particular the formal logic, which adopts specialist formal language to demonstrate the logical truth of composite statements, determining their truth by the component atomic statements.

I observe again that the truth or falseness of a statement implies that we can interpret it in a known language. Personally I cannot affirm whether the proposition "Пьеро является преподавателем" (Peter is a teacher) is *true* or *false* since I cannot understand its meaning, as it is given in an unfamiliar language; neither am I able to recognize the *signifier* of signs nor identify a *signified* and thus determine any *descriptions* and *definition*.

I shall conclude by reminding the reader that the truth or falseness of a statement (concerning what we have perceived and translated into linguistic signs) depends on the sensory organs of a person "A" who must interpret that proposition and compare it to the states of the world or with other propositions on these states. This implies that a proposition by this person "A," characterized by a given set of sensory organs, affirming the truth of a des(o) or a def(o*) may not be verifiable by another individual, "B," with a different set of sensory organs, unless the latter can "translate" the statements by "A" into a language that constructs statements that can be verified by "B." An example of this is the quote below which, though referring to humans, can be extended to any communication context involving subjects equipped with different sensors.

> The linguist manipulates the syntactic, phonological, and semantic forms and judges and/or asks native speakers to judge whether the consequences are a well-formed sentence in the language, an ambiguous string or any one of an array of numerous other possibilities. The relevant reference point by the very nature of the research is internal to the bearer of the internal grammar—the native speaker himself. To put the matter in a somewhat different form, suppose that we succeeded in constructing an instrument that purportedly arrived at the same judgments for visual inputs as those possessed by normally sighted people.

> How would we know whether the instrument worked? The answer clearly is that we would accept the instrument as accurate if and only if the responses of the instrument matched those of normally sighted people. In other words, we would calibrate the instrument by using precisely the same set of judgments (intuitions) reported by the people involved that we presently use in the absence of such an instrument. Thus in fields where the patterning under scrutiny is patterning of the behavior of human beings, the reference point and the source of the judgments will necessarily be the human being (Bostic St. Clair and Grinder 2001, p. 76).

9.13 Scientific Laws and Theories as Definitions and Conjectural Models of the World

Bateson's model can be extended and operationalized to also take in the important topic of *scientific explanations*. In fact, human thought is not limited to the identification and description of objects and the positioning of technical definitions. Instead such operations represent the premise for more advanced mental activities through which the conscious (and intelligent) cognitive system succeeds in giving explanations in order to understand certain phenomena in the world and to forecast these, in order, where possible, to control events.

"Advanced" mental activities are those directed contemporaneously at a group of objects, examples of which are:

(A) SET FORMATION: mental activity whereby different objects are considered as belonging to a set, with regard to specific D_n (books with a cover), d_n (books with a red cover), or Δd_n (books by Italian or English authors). "Sets" are formed as objects resulting from set formation, which itself appears to be an "essential" mental activity: I do not describe the object "set" but I form it as an independent object of observation composed of *simple, separate objects*: that is, the elements of the set. Sets are mental objects resulting from the application of a technical definition to distinct objects:

$$[\text{SET}] = \left\{ \left[\text{des O}^1\right] \cup \left[\text{des O}^2\right] \cup, \ldots, \cup \left[\text{des O}^M\right] \right\}.$$

(B) CLASSIFICATION: this is the operation of forming sets (or that of technical definition) applied several times to form subsets, or classes, of objects of observation. Each class is characterized by one or more dimensions and/or one or more ranges of admissible variation. More precisely, each class is characterized by a specific "technical definition for the class." The operation of "identification" (recognition) is associated with classification (understanding).

(C) SYSTEMATIZATION: the activity that takes place, as we know, when we observe objects linked by certain relations, with the result that these objects are interrelated and lose their observational individuality, becoming elements in a new entity: the "system"; a system is a composite object formed by *simple, unified objects* including a network of relations:

$$[\text{SYSTEM}] = \left\{ \left[\text{des O}^1\right] \cap \left[\text{des O}^2\right] \cap, \ldots, \cap \left[\text{des O}^M\right] \right\}.$$

(D) LEGALIZATION: the activity which searches for and identifies, in the observed universe, various types of regularity among the objects of defined sets. Various types of observation repeated with *regularity*, such as *technical descriptions* of events, phenomena and objects, including their dimensional states, even under different observational conditions, lead the "mind-processor" to *legalization*—that is, to *generalizing analogical abstraction*—which results in those regularities; these regularities are *defined* as *empirical laws* and constitute valid models of knowledge for all the observed objects. Every "law" derives from an *inductive* process; that is, from an *abstraction* whereby the mind *generalizes* for "all the objects x" (in formal logic we write (x), or $\forall x$):

- *Specific dimensions* (all flying animals must have wind sensors)
- *Specific dimensional states* (all arachnids have eight legs)
- *Specific relations* (all blue-winged butterflies have antennae as long as their forewings)

The mind observes "a finite number, N, of objects x"; the inductive generalization expressing a relation R between two classes of objects, x_n and y_n, can be expressed, for example, as follows:

$$
\begin{array}{ll}
\text{IF} & [x_n R y_n] = \text{"TRUE" for } n = 1, 2, \ldots, N \text{ (observation)} \\
\text{THEN} & (x)(y)[x \, R \, y] = \text{"TRUE"} \quad \text{(inductive generalization)}
\end{array}
\tag{9.15}
$$

(9.14) leads the mind to make the following deduction:

$$
\begin{array}{ll}
\text{IF} & (x)(y)[x \, R \, y] = \text{"TRUE"} \\
\text{THEN} & x_{N+1} R \, y_{N+1} = \text{"TRUE"}
\end{array}
\tag{9.16}
$$

The validity of this inference process is based on a number N of specific observations held to be sufficiently adequate to make the generalization plausible. Each new observation which confirms the relation (9.15) observed until that moment "validates" the generalization. Only a single contrary observation is needed to reject the generalization, unless it is transformed into a probabilistic generalization based on frequencies. In its most elementary form:

$$
\begin{array}{ll}
\text{IF} & [x_n R \, y_n] = \text{"TRUE" for } n = 1, 2, \ldots, M < N, \\
\text{THEN} & (x)(y)[x \, R \, y] = \text{"TRUE" with probability } (M/N)
\end{array}
\tag{9.17}
$$

When a scientist says that the evidence is sufficiently good to establish a scientific hypothesis, he means that the evidence is such that the hypothesis could be inferred from it according to one of the recognized principles of scientific inference. His interest is in the adequacy of the evidence for this purpose, for this is what would be in dispute if a fellow scientist questioned the 'well-groundedness' of the hypothesis. [...].

Thinking only of the considerations which are relevant to a scientist, a philosopher will be quite right in taking the question as to whether or not an inductive inference is valid and the belief in a scientific hypothesis yielded by it is a reasonable belief as being a question as to whether or not the empirical evidence is such that the hypothesis can be obtained from it by following recognized 'rules of procedure' (Braithwaite 1953, p. 262).

For an *empirical law* (*norm* or *generalization*) to be defined as *scientific* it must:

1. Be presented as the following equivalent type of statement (9.15): "if [des A] then [des B], always"; $(x)(y)([des x] \rightarrow [des y])$; $\forall[des A]$, then we observe [des B], always; [des B] does not exist without [des A], etc.
2. Have empirical content; without empirical content it can at most be a formal law.
3. Present relations between objects belonging to *open sets* that are connotatively defined and whose extension is not finite or entirely known; this is thus valid not only for observed objects but also for all those objects that have the connotations that define a set, even if they have not yet been observed.
4. Not present a relation that derives from conventions or the application of a procedure.
5. Be *verifiable* or *falsifiable*; that is, confirmed by favorable cases, or positive examples, or confuted by unfavorable cases, or contrary evidence.

6. Be *coherent* with other accepted scientific laws and permit *deductions* when included in deductive argumentation.

(E) THEORIZATION: the activity of an evolved mind whose aim is to *interpret* and *explain* the perceived reality, even going beyond known scientific laws. Theorization is needed to give "closure" to scientific explanation (Sect. 9.15.2). Through this activity scientific theories are formed; that is, formalized systems of coherent hypotheses or conjectures, linked together, which aim to justify, and possibly explain, the existence of "laws" and their connections, thus forming a unitary descriptive and predictive corpus. Having identified regularities and *laws* for a given observed universe, the observer/mind then tries to understand the *reason* for their existence. He thus presents *theories*—that is, hypotheses or conjectures (systems of hypotheses)—that could justify the affirmed regularities. The theories (as well as the single laws) can be interpreted as *formal hypothetical definitions* of the *observed universe* presented by man to completely *describe* that universe. Theories, as *explanatory hypotheses*, must not only contribute to *explaining observed* facts but also permit *forecasts* about *observable* facts. In their so-called realistic conception of knowledge, Stephen Hawking and Leonard Mlodinow, recognize that all we can know about "reality" consists of networks of world pictures, or general models, expressed even through mathematical language; only the systematic connection of a set of observations accompanied by a conceptual model and by rules connecting the concepts of the model to the observations enables the mind to explain those observations by connecting them by rules to concepts defined in models.

> In the history of science we have discovered a sequence of better and better theories or models, from Plato to the classical theory of Newton to modern quantum theories. It is natural to ask: Will this sequence eventually reach an end point, an ultimate theory of the universe, that will include all forces and predict every observation we can make, or will we continue forever finding better theories, but never one that cannot be improved upon? We do not yet have a definitive answer to this question . . . (Hawking and Mlodinow 2010, p. 8).

Multiple, equally valid, world pictures exist; therefore, science requires multiple models to encompass existing observations:

> Like the overlapping maps in a Mercator projection, where the ranges of different versions overlap, they predict the same phenomena. But just as there is no flat map that is a good representation of the earth's entire surface, there is no single theory that is a good representation of observations in all situations (Hawking and Mlodinow 2010, p. 9).

9.14   Intelligent Cognitive Systems: The Power of Models

Having demonstrated Bateson's hypotheses that the mind is nothing other than a perceiver, calculator, and analyzer of various levels of differences, we can now define a *conscious* cognitive system as an *autopoietic system* that through the "mind" is able to distinguish differences as external disturbances and develop knowledge by adapting to these disturbances, thus forming concepts that allow

the mind to link its own mental representations to form an "external world" to which it has no direct access.

Following Bateson, the "mind" thus represents the computational system capable of constructing knowledge and transforming the autopoietic system into an *observer system* that can *describe* objects and produce *definitions* of concepts from which a *knowledge process* can be derived.

> What we know is generally considered to be the result of our exploration and understanding of the real world, of the way things really are. After all, common sense suggests that this objective reality can be discovered. . . .How we know is a far more vexing problem. To solve it, the mind needs to step outside itself, so to speak, and observe itself at work; for at this point we are no longer faced with facts that apparently exist independently of us in the outside world, but with mental processes whose nature is not at all self-evident (Watzlawick 1984, editor's Foreword).

> The world, as we receive it, is our own invention (von Foerster 1984, online p. 45).

I shall define a *cognitive system* as *intelligent* when, through the "mind," it not only "constructs" the world but also produces a codification process of the differences-information, creating a system of *semes*, a *language* through which the *cognitive system* links up with other *cognitive systems* in *formal communication processes* that allows it to form *scientific laws* and develop *argumentations*. It thus follows that a necessary and sufficient condition for a *cognitive system* to also be *intelligent* is that at the same time it is *autopoietic* and capable of developing formal *communications behavior* with other systems it is linked to. In the end, this is the ultimate meaning of Turing's Test (1950), which we are all familiar with. This is the basic hypothesis of *functionalism*, one of the most important research fields in cognitive science. Extrapolating from Andy Clark's view of the theory of symbols (signs possessing conventional meaning):

> The functionalist is in many ways the natural bedfellow of the proponent of the physical-symbol-system hypothesis. For the physical-symbol-system hypothesis claims that what is essential to intelligence and thought is a certain capacity to manipulate symbols (Clark 1989, p. 19)

> A *physical symbol system*, according to Newell and Simon (1976, pp. 40–42) is any member of a general class of physically realizable systems meeting the following conditions:

> 1. It contains a set of symbols, which are physical patterns that can be strung together to yield a structure (or expression).
> 2. It contains a multitude of such symbol structures and a set of processes that operate on them (creating, modifying, reproducing and destroying them according to instructions, themselves coded as symbol structures).
> 3. It is located in a wider world of real objects and may be related to that world by designation (in which the behavior of the system affects or is otherwise consistently related to the behavior or state of the object) or interpretation (in which expressions in the system designate a process, and when the expression occurs, the system is able to carry out the process).

> In effect, a physical symbol system is any system in which suitably manipulable tokens can be assigned arbitrary meanings and, by means of careful programming, can be relied on to behave in ways consistent (to some specified degree) with this projected semantic content (Clark 1989, pp. 4–5).

With our powerful minds and our capacity to construct instruments to expand the limits of our sensors and effectors, we humans are able to expand our intelligence by constructing *descriptions* and *definitions* relative to ever wider "portions of the world," by observing and understanding the "increasingly smaller," the "increasingly larger" and the "increasingly more distant." The phenomena observed are many, interconnected, variable and multidimensional, giving rise to all types of differences. Bateson's mind finds it difficult, if not impossible, to mentally describe such a vast "world," and as a result it is more and more difficult to communicate to other persons the results of our observations; ordinary language, spoken or written, is not always adequate to form and efficiently transmit complex representations, which is why man has always sought to define specific signs to construct efficient forms of representation, even in alternative languages. The signs and languages of geometry, logic, and mathematics greatly expand the mind's capacity for synthesis and analysis. Moreover, thanks to the continual advancements in these languages man has succeeded in understanding (defining) the unobservable (nondescribable) objects: from atoms to gravity, genes to evolution, complex dynamics to chaos, etc.

In short, the human mind can efficiently form descriptions and concepts by using appropriate *signs* and *operators*, defined in a specific manner, in order to construct increasingly more efficient models of the observable universe. For example, all the models in this book allow us to define the control process using the graphical tools I have adopted; however, the use of differential equations and integral calculus, or the use of Powersim, MatLab, and so on, would add other descriptions and definitions. And then there is the written language of music which, with its unique notation, allows us to codify the sounds-notes (primary differences), their pitch, intensity, timbre (higher-order dimensions) in order to form musical phrases which together form well-structured models (sonatas and symphonies, for example) and to transmit these in space and time so that they can be newly translated into sounds by reconstructing in the minds of future musicians and listeners the same model as that of the composer. If today we can listen a thousand times to Beethoven's 9th symphony, this is thanks to the system of musical signs that allowed the composer to leave behind his musical score, which a thousand interpreters can now newly translate into sounds.

In fact, the universes of observation, being composed of a group of objects, all of which multidimensional and often closely interrelated, are almost always too complex to lend themselves to observation and comprehension. Therefore, observational activities could not effectively and efficiently be undertaken if the observed did not formalize the results of such observations using appropriate *meaningful semic representations* that constitute *models*. As such, models are "depictions" of the results of observational activity and are instrumental in efficiently communicating the observational results for reasoning, argumentation, and explanation (Sect. 9.15.2).

We can distinguish between *mental models* and *formal models*. The former are the models the *intelligent* cognitive system creates in its "private" mental sphere to represent the content of the meaning it wishes to transmit (if it is the sender) or that it must interpret (if it is the recipient), which it uses to evaluate "the world" and

make instinctive decisions. *Formal models* are made up of systems of concrete signs, linguistic or symbolic, which formalize and transfer a meaning and are used to communicate and enable weighty decisions to be made (Sect. 9.15.1).

Though approximate, *mental models* indispensable for *intelligent* thought, as they are the basis for an understanding of reality. The psychologist Philip Johnson-Laird analyzed their cognitive function in his book *Mental Models: Towards a Cognitive Science of Language, Inference and Consciousness* (1983), an analysis that was taken up again in many of his subsequent works.

> The psychological core of understanding, I shall assume, consists of having a "working model" of the phenomenon in your mind. If you understand inflation, a mathematical proof, the way a computer works, DNA, divorce, then you have a mental representation that serves as a model of an entity in much the same way as, say, a clock functions as a model of the earth's rotation . . . Many of the models in people's minds are little more than high-grade simulations, but they are none the less useful provided that the picture is accurate (Johnson-Laird 1983, pp. 2, 4).

The following excerpt, which comes from one of his recent works, lets us see how the mental models make *thoughts* revealed by some linguistic expression corresponding to simplified *mental representations*.

> [The sentence] a triangle on the right of a circle has a single mental model, which we depict in this diagram:

> The left-to-right axis of the model corresponds to the left-to-right axis of a scene, and the disposition of the triangle and circle in the model corresponds to their disposition in a scene for which the assertion is true. The model represents an indefinite number of possibilities that have in common only that a triangle is on the right of a circle. Of course, the relative sizes of the figures in the model, their distance apart, and so on, play no role in reasoning from the model [. . .] (Johnson-Laird and Khemlani 2013, p. 6).

Peter Senge, in The Fifth Discipline, deals with mental models on more than one occasion, offering these definitions:

> "Mental models" are deeply ingrained assumptions, generalizations, or even pictures or images that influence how we understand the world and how we take action. Very often, we are not consciously aware of our mental models or the effects they have on our behavior. For example, we may notice that a co-worker dresses elegantly, and say to ourselves, "She's a country club person." About someone who dresses shabbily, we may feel, "He doesn't care about what others think" (Senge 1990, p. 8).

> Our 'mental models' determine not only how we make sense of the world, but how we take action. Mental models can be simple generalisations such as 'people are untrustworthy' or they can be complex theories, such as my assumptions about why members of my family interact as they do. But what is important to grasp is that mental models are active—they shape how we act. If we believe people are untrustworthy, we act differently than if we believed they were trustworthy" [. . .] Why are mental models so powerful in affecting what we do? In part, because they affect what we see. Two people with different mental models can observe the same event and describe it differently, because they've looked at different details (Senge 1990, p. 160).

Jay Forrester pointed out this "mental model of a mental model":

> The mental image of world around us that we carry in our heads is a model. One does not have a city or a government, or a country in his head. He has only selected concepts and relationships, which he uses to represent the real system (Forrester 1971, p. 213).

Mental models are the result of, and at the same time the basis for, the cognitive activity we are continuously producing and represent the primary form of representation of the world and the relations we observe in it. Nevertheless, mental models, precisely because they are produced "in the private sphere," have a clear limit: not only can they be imprecise and vague but also they are often *erroneous* and, if not corrected through subsequent verifications or educational processes, can produce seriously harmful results. There are an endless number of examples of this, some of which it is useful to present here.

Before Isaac Newton the prevailing mental model for gravity was, in simple words, that: "The apple breaking off from the branch falls toward the ground, attracted by the earth, following a trajectory directed toward the center of the globe." Newton changed the mental model of gravity as follows: "Apples are attracted to the earth as much as the earth is attracted to apples, following a trajectory determined by their mutual centers of gravity." Another prevailing mental model was: "All bodies in motion are destined to stop, unless a force intervenes to keep them in motion." Newton's First Law of Motion changed this mental model by introducing the principle of inertia: All bodies will continue moving in a uniform straight line (or remain still) unless some other force causes it to change this state." Another obvious erroneous mental model, which, unfortunately, is widespread even today, even among educated persons, is the following: "When you have a cold your nose runs; I see that your nose is running, so I must stay away from you because I don't want to catch your cold." The logical mistake is clear as soon as we consider that my nose could be running because I have just finished crying my eyes out or because I suffer from hay fever.

Precisely to overcome these limits and risks due to reasoning and decisions based solely on *mental models*, man has always sought to accompany these, even at times replace them, with the more efficient *formal models* that allow him to represent and share, through specific language and using various types of signs and symbols, all thought content (Sect. 9.15.1), thereby allowing him to create science and improve the *explanatory* processes (Sect. 9.15.2).

Mental and *formal* models can represent objects of observation through their *descriptions* and technical *definitions*, as well as sets, systems, laws and theories, and entire universes of observation. The approximation to the cognitive "reality" that is to be represented is a necessary feature of every model, even operationally speaking, since it may be the only means of investigating a reality which is too complex, or of studying certain aspects of a single reality that cannot be manipulated and directly investigated. However, this approximation must not be overly vague but instead achieve maximum utility for an efficient and effective representation, communication, and understanding. If the approximation has such utility, then the model can be a substitute for the particular reality and the universe of

observation in further observational research. The differences and dimensions which the mind identifies or constructs in the model become direct objects of the observational activity; the signs of the model replace the mental objects in the processes regarding knowledge and recognition, the formation of laws, and the construction of hypotheses.

I shall conclude by going back to Sect. 1.2. I am convinced that intelligence depends on the ability to construct coherent and sensible models in order to acknowledge and understand the "world." I believe that:

> If the capacity to *see* and not simply look at [the world] depends on the ability to construct models to understand, explain and simulate the world, then the most useful and effective models to strengthen our intelligence are the systems ones based on the logic of Systems Thinking (Mella 2012, p. 6),

since these models allow us to represent complexity, dynamics, and change (see Sect. 1.5).

9.15   Complementary Material

9.15.1   Models

It is useful to mention several important types of models, emphasizing that it is possible to combine these types to obtain highly structured classes of models, such as:

(a) Models in lexical or specialist languages
(b) Iconic and formal models
(c) Models of objects and relations
(d) Concrete and abstract models
(e) Descriptive and operational models
(f) Scale and analogic models
(g) Static and dynamic models
(h) Qualitative and quantitative models
(i) Deterministic and stochastic models

Models in lexical or specialist language. Lexical language (English, for example) represents the basic instrument for translating thought content into transmissible models. If I have to describe my sister's wedding to a friend I have to use lexical language, transforming the mental *descriptions* of the ceremony into a flow of words following an informational path of my choosing (I first describe the setting, the spouses or the guests, etc.). On the other hand, if I must represent string theory, then lexical language is not suitable and I will have to use the same mathematical language theoretical physicists use to construct their models on the theory of the universe.

Models of objects and models of relations. Here the distinction concerns the thought content to which we give priority in representing the models. Models of objects depict the objects of observation or individual *mental models*, simple or composite (a firm, landscape, etc.); models of relations instead represent the relations that bind together several objects, or the components of a structured object. Models of sets and systems belong to this class of models.

Concrete and abstract models. This dichotomy refers to "material" the model is composed of. Concrete models are made up of materials whose properties concretely represent the dimensions of the objects they depict (statue, scale model of a machine, plastic in a building structure, etc.); abstract models instead represent the objects through signs (words in a language, signs and symbols, designs, graphs, etc.).

Iconic and formal models. This classification regards the "form" of representation of the original. Iconic models are constructed so as to "resemble" the original (photograph, films, a square, a design, a wax statue of Napoleon, a reduced-scale airplane model with the same colors as the original, and so on). Formal models instead represent the original in "schematic" terms and do not allow us to recognize the original thing being represented, though we can perceive the dimensions intended to be represented.

Within *formal* models we can distinguish among: *verbal, symbolic, logical-mathematical,* and *schematic* models. *Verbal* models are expressed using an everyday language, such as in the reasoning process: "If all men are mortal and Greeks are men, then Greeks are mortal." *Symbolic* models instead employ a symbolism based on an agreed upon and recognizable code. *Logical-mathematical models* rely on the symbolism typical of the language of logic and mathematics. For example, let us assume we agree on the following use of symbols: if H stands for Men, G for a Greek man, m for a mortal, x for a variable object, then: "any x which belongs to H also belongs to m; any x that belongs to G also belongs to H; any x that belongs to G must be an 'm'." Obviously to express models of dynamic and Control Systems, *schematic* models may be more appropriate, since these immediately depict the object to represent through graphical signs and not linguistic ones in the strict sense of the term. For dynamic systems the language of differential equations or difference equations could be more appropriate, accompanied by graphical representations of dynamic systems.

The advantage of *iconic* models is that they are able to immediately perceive the dimensions of shape, color, weight, etc., which often cannot be easily expressed in formal terms: "What color was the bride's bouquet?" To answer this question a photograph would be a better iconic model. Abstract models, though possessing the disadvantage of not being able to perceive similarities, have the advantage of being able to perceive analogies and commonly accepted conventions, and they are more easily constructed and transmitted. More specifically, linguistic models, literary, logical or mathematical, allow for a nearly unlimited representation of the depicted objects. Schematic models, though having a more limited range of possible representations, offer the advantage of immediacy and the possibility of highlighting relations in structural terms. Let us suppose we wish to represent the model of a system expressing the spatial positioning of 5 objects: A, B, C, D, and E. A verbal
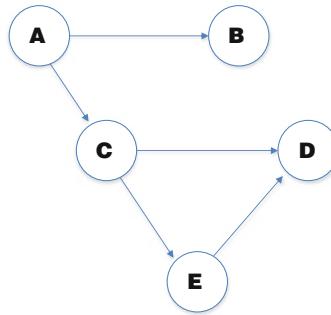
Fig. 9.13 Schematic model of interrelated objects

language would allow us to describe the system, for example, using the following annotation: "A is to the left of B; A and B are above C and D; C is to the left of D; E is below C and D; A is linked to B and C; C is also linked to D and E; E is also interrelated with D." Thanks to the use of the signs "left," "above," "below," etc., which order our mind to undertake the mental operations for perceiving the arrangement, our mind can form a mental model, perhaps with some difficulty, given the number of objects and the interrelationships. The schematic model in Fig. 9.13, where the arrows indicate relations and the positions are visually apparent, is much simpler and more significant. Note that almost all the models of systems presented in this book are schematic.

Descriptive and operational models. These classifications are determined by the operational objectives the subject wishes to achieve through the model. Descriptive models serve to describe or explain the object they represent by presenting the relevant dimensions through some language or schema; they can represent a technical description (a red-sided triangular figure; a green surface; a 5 cm base; 4 cm in height; 42° vertex angle, etc.) or a technical definition (if AB, BC, and CD are linked segments forming a figure lying in a plane, then ABC is a triangle). *Operational* models are structured to simulate the "behavior" or dynamics of the phenomenon being represented (a mechanism to simulate earthquakes; systems of differential equations to simulate growth in an economic system; Excel programs to simulate the dynamics of Control Systems, etc.).

Scale and analogic models. These classifications are based on the type of approximation the model introduces in the represented phenomena. In *scale* models all the dimensions of the model are maintained, but some dimensional measurements are presented with a different scale through enlargement or reduction. Analogic models, on the other hand, modify the observed dimensions and represent the thought object with dimensions having similar features. In other words, while scale models maintain the initial quantitative dimensions and the ratios between them (changing only the distance), analogic ones depict the original dimensions through different dimensions while respecting the ratios (and/or relations and/or distances) already observed in the represented object. Since the construction of

scale models implies the conversion of the measurements based on scale, which changes the units of measurement with respect to those used in quantifying the original objects, it is necessary to avoid the mistake of considering the original dimensions as all being equally and arbitrarily modifiable. In fact, some dimensions can be reduced in scale only up to certain levels; for example, the surface tension of water molecules can have no influence on a model of a ship with a scale of 1:100, but it would undoubtedly be relevant in cases of greater reductions in scale: a needle will float on water, but a needle whose weight is increased tenfold will inevitably sink. Even with the construction of *analogic* models it is necessary to proceed carefully. For example, if it is not possible to simulate the flight of an airplane by constructing a model that can actually fly, we can build a similar and operational model by simulating the flight in a wind tunnel.

Static and dynamic models. The former depict structural relations among the components of a system that do not evolve over time (a company's balance sheet is a typical static model of values). Dynamic models, on the other hand, consider time (or space) as a fundamental variable; every dynamic dimension of the original must be appropriately depicted in the model (for example, models that represent self-controlled systems are dynamic).

Qualitative and quantitative models. The distinction here refers to the type of dimensions held to be fundamental more so than to the results achieved. Models that, irrespective of their conformation, represent quantitative dimensions from the original are *quantitative*. Qualitative models instead highlight the qualitative dimensions. The schematic model in Fig. 9.13 is typically qualitative, since it emphasizes the *name* of the elements and the order relations among them.

Deterministic and stochastic models. In this case the distinction regards the way the relations represented in the model are determined, typically when the model concerns dynamic systems. Probabilistic models hypothesize that the inputs and/or outputs are random variables whose probability distribution is assumed known.

9.15.2   The Explanation Process and Its Operational Closure

When the *mind* of man is faced with phenomena or events he cannot *understand*, a situation of *ignorance* exists that gives rise to *curiosity*, which leads to a *question* that requires an *explanation*; that is, guidance toward the construction of a model from which those phenomena or events could derive. Technically speaking, *explaining* a phenomenon F (*explanandum*) means *constructing a model for it based on some other knowledge*, thereby identifying a *succession* of statements; that is, constructing a deductive *argumentation* (*explanans*) from which it is possible to derive F as a *valid* conclusion (Hempel 1965). Every *explanandum* requires an appropriate *explanans*. Knowledge progresses due not only to new observations and mental models but also to subsequent explanations beginning

from models that are well known or assumed to be so. Nevertheless, explanation is a *backwards-recursive process*: every *explanans* of an explanation becomes, in turn, an *explanandum* that can require a subsequent explanation, and so on up the line, just like the "game of *why*" that all children love to play. Any ignorance is explained with some piece of knowledge; but this, in turn, is based on other knowledge (Keil 2005).

I indicate with $E(n)$ the *explanans* and with F the *explanandum*; the recursive explanation then has the following structure:

F because $E(n)$,
$E(n)$ because $E(n-1)$,
$E(n-1)$ because $E(n-2)$,
...
$E(0)$ End.

The *explanation process* is *closed* when we reach $E(0)$; that is, an explanans with no further explanandum. I define $E(0)$ as the *operational closure* of the explanation, which represents the *point of ignorance*, the limit to knowledge. $E(0)$ can denote two possible forms of ignorance:

$E(0)$ = "That which is not yet known." This represents *reversible* ignorance, which can take on the form of: (1) *temporary ignorance* ("research in progress ...," "we are checking," etc.): (2) *disinterest*, that is, satisfaction with $E(1)$ ("I'm not interested," "Enough already!"); (3) *secretiveness*, that is, the desire to hide knowledge ("I can't explain it to you").

$E(0)$ = "That which we can never know." In this situation, *operational closure* represents *irreversible ignorance*, which can take various forms: (1) *permanent ignorance*, that is, that which *reveals itself* incapable of *ever* being known (Heisenberg's indetermination principle, Gödel and Turing's theorems, and others that are similar); (2) the *mysterious*, that is, the *unknowable* (ineluctable principles, myths, God, metaphysics or agnosticism, acts of faith, etc.); (3) the *postulate*, that is, a nondemonstrable, only assumed, origin from which to derive the explanation ("We hypothesize that ...," "Given that ..."; "We assume that ...," etc.); (4) *chance*, in all its manifestations (quantum randomness, genetic mutations, initial impulses of combinatory systems, etc.); (5) *necessity*, that is, the self-sufficiency of $E(0)$ ("It must be like this," "It can't be otherwise," "The immobile engine," etc.). In these cases operational closure is an *explanatory dogma*.

In very concise terms, an *explanation* is:

(a) *Complete*, if it also includes its own operational closure; otherwise it is *partial*.
(b) *Global*, if its operational closure is a *dogma* and does not allow for further explanations; otherwise it is *intermediate*.
(c) *Coherent*, if the explanans does not include elements that contradict other explanations.
(d) *Valid*, if it derives from a correct inference.
(e) *Truthful*, if the $E(t)$ do not contain false elements that have nevertheless been declared to be true.

(f) *Logical,* if it allows us to derive the explanandum through a satisfactory operational closure on the basis of a certain stock of knowledge and in a given social context.
(g) *Efficient*, if the explanans is composed of a minimum number of premises.
(h) *Effective,* if it is complete, logical and allows for verification or forecasts.

The coherence, validity, and truthfulness of an explanation are, according to the laws of science, always verifiable; logic, efficiency, and effectiveness are instead subjective. Every attempt at a global explanation is based on irreversible ignorance. After having outlined the main features of the explanation process, we will examine five types of explanation processes that I consider the most relevant and which rely on different premises.

Common Sense, Descriptive, or Contingent Explanations

The so-called common (simplistic, unsophisticated, or descriptive) explanation is the one we daily apply to justify some phenomenon or to answer some question. In the common explanation the phenomenon or question to explain is derived from models that usually describe what we must explain without looking for the under-lying scientific causes; "the light has come on because Aldo has flipped the switch"; "he got a fever because he was bit by an insect," etc. The common explanation at times relies on intentions ("the light has come on because Aldo had to go into the garage"), chance ("I met Aldo because he happened to be passing by"), experience ("it's raining because I went out without my umbrella"), regularity or common sense laws ("this is a rainy spring because the winter was dry"). The same phenomenon can thus give rise to a number of common explanations that depend on the most convincing factors and that usually allow for one or at most two levels ("you made a mistake because you were distracted. What are you thinking of?"; "I'm thinking of what I'm going to do tonight"). This type of explanation is usually expressed descriptively and does not make use of sophisticated qualitative models (Hempel, 1965). Included in this typology are historical "explanations" as well as "attribution theory," in order to take into account one's own and others' behaviour, attributing this to the most likely "causes," such as justification (Heider, 1958). Another theory included in this typology is the 5-Ws—Who, What, When, Where, and Why—which accounts for most of the explanatory needs in our daily lives. The general structure of the common explanation is as follows:

Explanans	P = descriptive, contingent or simplistic, unsophisticated Premises: – functions, – aims, – intentions, – causes, – chance, etc.
Explanandum	F = phenomenon to explain
Conclusion	F occurs because the premises P have occurred
Operational closure	This is the last of the premises that recursively explain F

The contingent explanations can easily become tautological—and thus nonexplanatory—when one tries to explain a phenomenon simply through its

definition: "Inflation exists because all prices are rising"; "internationalization is an inevitable process because socio-political barriers are overcome," etc.

Classical Scientific Explanations

The classical explanations are those that are used in any scientific context where the explanans is a model which, in addition to the *initial conditions* of the phenomenon to explain, C, also includes the *causal laws* or functional relations, L, as well as the *scientific theories* or assumptions that can take into account the explanandum, according to the following schema:

Explanans	T = scientific theories and postulates
	L = scientific laws or statistical generalizations
	C = initial conditions
Explanandum	F = phenomenon to explain
Conclusion	F is observed because, given the initial conditions C, it follows from the laws L, if we accept theory T
Operational closure	The fundamental theories and postulates represent the operational closure of the explanation

By varying the theoretical context and/or the laws adopted and/or the initial conditions, we also vary the explanation (Cupples 1977). This form of explanation has become recognized as the scientific explanation *par excellence*.

> Forty years ago a remarkable event occurred. Carl G. Hempel and Paul Oppenheim published an essay, "Studies in the Logic of Explanation," which was truly epoch-making. It set out, with unprecedented precision and clarity, a characterization of one kind of deductive argument that, according to their account, does constitute a legitimate type of scientific explanation. It came later to be known as the deductive-nomological model. This 1948 article provided the foundation for the old consensus on the nature of scientific explanation that reached its height in the 1960s (Salmon 1990, 3).

> Hempel claimed that "there are two types of explanation, what he called 'deductive-nomological' (DN) and 'inductive-statistical' (IS) respectively." Both IS and DN arguments have the same structure. Their premises each contain statements of two types: (1) initial conditions C, and (2) law-like generalizations L. In each, the conclusion is the event E to be explained [...] The only difference between the two is that the laws in a DN explanation are universal generalizations, whereas the laws in IS explanations have the form of statistical generalizations (IEP online, 2-3).

Procedural Scientific Explanations

The procedural is a very common type of explanation, even though it has not received proper attention in the literature. It is used whenever a phenomenon does not derive from a particular model but rather appears to be the result of some elaboration or calculus, or the application of some algorithm, procedure, or program (Gibbon 1998). When we ask ourselves why the solution to extracting the

square root of an expression does not correspond to the answer in our textbook, why our ticket wasn't drawn in the lottery, or why we ran into the bumper of the car in front of us, we must look for the answer in the procedure followed for the calculation or lottery drawing, or we must examine our parking attempt. The procedural explanation can take on the following structure.

Explanans	P = procedure whose application produces F
	C = conditions for applying the procedure
Explanandum	F = phenomenon to explain
Conclusion	F derives from P applied under the conditions C
Operational closure	The procedure represents the operational closure of the explanation

The procedural explanation comes under the generalized heading of systemic explanation.

Systemic Scientific Explanations

The systemic explanation must be used when the explanandum cannot be reduced to a model that includes laws and theories (classical explanation) or results from the application of a procedure; instead it must be considered a phenomenon connected to the dynamics of some system process whose model we are trying to uncover. The systemic explanation must highlight the processes P that have generated F; the systemic structure S that supports those processes; the programs Π that guide the latter; and the environment E that conditions them (Mella 1997, Sect. 2.16). The model has the following structure:

Explanans	S = systemic structure that generates the processes that give rise to explanandum F
	Π = programs that make the structure generate the processes
	P = processes generated by the structure S through the operations of the programs Π
	E = environment to which the system is coupled
Explanandum	F = phenomenon to explain
Conclusion	F derives from P produced by S in E by means of Π
Operational closure	The system (environment-structure-programs-processes) represents the operational closure of the explanation

The systemic explanation is thus more powerful, as it can take into account and justify any phenomenon whatsoever, from global warming to population dynamics, the deviation of the route of a space probe to the spread of epidemics. Most of the explanations presented in this book by using the Control Systems model belong to the class of systems explanations. This class is not an alternative to other explanations but rather encompasses them, becoming a type of even more general explanation.

Teleological Explanation

Systems explanations can be kept separate from *teleological explanations*, which try to take into account the behavior of a system (usually a biological one) by using the notion of goal (Lennox 1992). Such explanations regard systems with objective, or "goal-directed" systems to which are applied the Control Systems presented throughout the book.

> The necessity of teleological explanation for biology has been questioned by the formalist tradition. We have seen that the ideal of explanation for the formalists is the deductive model. Many attempts have been made to incorporate teleological explanation into the deductive model. Indeed such an incorporation is necessary if there is to be a formal reduction of biology to physics and chemistry. The basic line of argumentation is to subsume biological "goal-directed" systems under the wider category of "directively organized" systems which can apply to both living and inorganic systems (Plamondon 1979, p. 153).

Teleological explanations are evolutionary explanations that justify certain biological phenomena linked to individuals and populations and which derive from some necessary aim: a type of *backward causation*. As an example of a teleological explanation, let us assume we wish to explain "why the winter coat of the hare becomes white." The evolutionary teleological explanation could be presented as follows:

1. Differences in color exist in all natural hare populations.
2. If in winter a hare has a white coat, it is better able to camouflage itself and improve its chances of survival.
3. The greater its chances of survival are, the greater are those that it will successfully procreate.
4. Thus, on average hares with a lighter-colored winter coat will have more offspring compared to those with darker coats.
5. The color of the winter coat is a hereditary feature.
6. Thus, the following generation will see a greater number of hares with a white winter coat.
7. After many generations, only hares with white winter coats will exist.

The model of the teleological explanation has the following structure:

Explanans	S = systemic structure that generates the processes that give rise to explanandum F
	G = Goal that S must achieve
	C = Control process carried out by S
	E = environment to which the system is coupled
Explanandum	F = dynamics of S to explain
Conclusion	F derives from S in E because C directs S toward G
Operational closure	The Control System directing S to G represents the operational closure of the explanation

The Best Explanation

While the classical explanation is particularly useful in the context of the experimental sciences, systemic explanation can also be used to explain individual, nonrepeatable events involving unique facts deriving from a system's behavior. The former bases its validity on the initial conditions of the phenomenon in question and on a system of laws and theories that represent the final closure of the explanation process. The systemic explanation is based on the search for the most appropriate system model for deriving the explanandum F and has as its operational closure the boundaries of the system and the extension and organization of the network of processes that constitute its functioning.

The best explanation is the one considered most likely; that is, the one that best satisfies the need for the individual or scientific community to take account of the explanandum. I do not believe anyone would turn to physiological or psychological laws and theories or to logistical processes to explain "why there is no more mayonnaise."

A good explanation must permit accurate forecasts. *Forecasting* phenomenon F means developing an argumentation which, based on laws and specific conditions, allows us to deduce F before we observe it (independently of that observation). It is immediately clear that the structure of the argumentation which produces the *explanations* or *forecasts* is entirely analogous; the difference is only in the time dimension. If the explanandum is "dated" (t_1) and the explanans (t_2), we have an explanation; if, on the other hand, the premises refer to the instant (t_1) and the conclusion to instant (t_2), the argumentation is in the nature of a forecast. While the excellence of the explanation depends on the extent to which the premises justify the conclusion, that of forecasts depends on the extent to which the conclusions are justified by the premises. Thus, scientific and systemic explanations are preferred since these argumentative structures are considered to be the best even in providing accurate forecasts of the occurrence of F.

In many circumstances a *realistic approach* must be followed; as Hawking and Mlodinow have stated (Sect. 9.13), it is necessary to constantly search for new forms of explanation that can explain the observations of facts, thus enriching in an increasingly coherent manner the networks of world pictures that can provide explanations and understanding. The explanatory process can thus be linked to the theory of Control Systems, since it develops dynamically, as shown in the model in Fig. 9.14.

When there are two equally likely explanations, the one whose explanans includes the theories, procedures, and systems the scientific community holds to be the simplest and most accredited is generally considered the better one. If a given scientific community is divided into groups that accept different operational closures (laws, theories, structures, procedures, programs, etc.), then alternative explanations for a given phenomenon are possible, all of which are equally accredited. The explanation becomes unequivocal only when progress in that scientific field leads to one operational closure prevailing over the others. Often those explanations
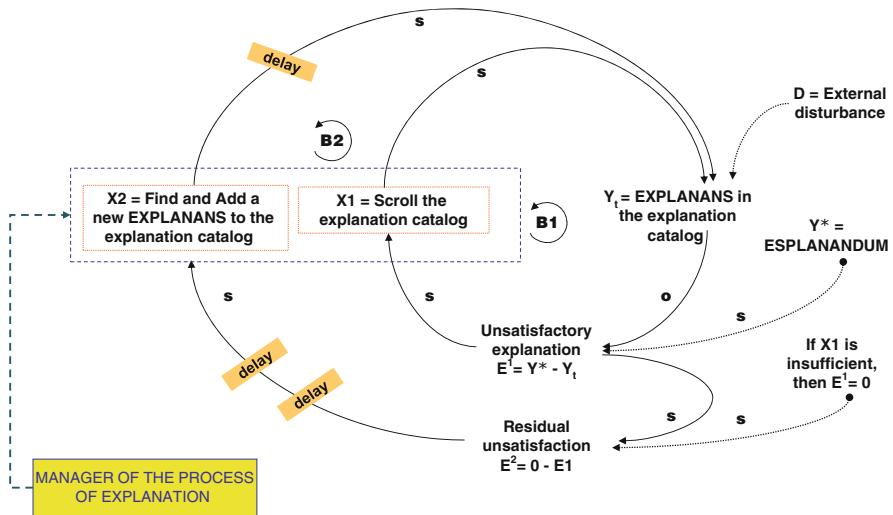
Fig. 9.14 The *Ring* representing the explanation process

are considered best that include among their premises only (or only a few) causal laws. However, I believe I have demonstrated that the inclusion of causal laws is often impossible when the phenomenon to explain derives from a system; in that case the systemic explanation is more convincing.

Finally, I would note that in order to understand and explain reality, Systems Thinking and System Dynamic approaches, jointly applied, allow us to identify the best systemic explanations by constructing models of dynamic, repetitive, and interconnected systems and simulating their operation (Mella, 2012). Objects observed from a static vision, nonrepetitive systems, individual phenomenon, simple causes, simple effects, a lack of memory: these are the errors Systems Thinking tries to eliminate.

9.16   Summary

To conclude this brief journey in the world of *Magic Rings*, this chapter has dealt with, even though with reference to a limited set of arguments, the important topic of knowledge in order to demonstrate that even thought processes, understood as emerging activities of the "mind" made possible by the processes of the brain, can take place only thanks to a large number of *Rings* operating at very high speed and in such a routine manner as to escape our attention. This chapter has focused on:

(a) Presenting Bateson's model of the "mind" as a calculator and transformer of differences (Sect. 9.2 and Figs. 9.1 and 9.2).
(b) Hypothesizing how the "mind" calculates similarities and analogies by constructing observational dimensions (Sect. 9.3).

(c) Indicating how the "mind" can describe objects (Sect. 9.4).

(d) Hypothesizing the process through which the mind compares objects through their technical descriptions (Sect. 9.5 and Fig. 9.4)).

(e) Presenting the *definition* process through which the mind generalizes and makes abstractions by moving from objects to *concepts* (Sect. 9.6 and Fig. 9.6).

(f) Constructing a model of a two-dimensional cognitive universe (Fig. 9.7) that represents the three paths of knowledge: *cognition of objects*, *understanding* concepts and *recognition* of objects (Fig. 9.8) as belonging to a particular concept (Sect. 9.7).

(g) Indicating the ways in which the human mind constructs meaningful technical definitions (Sect. 9.8).

(h) Hypothesizing the way in which the formal concepts of *cognition*, *understanding,* and *recognition* allow us to translate symbolic terms into "mental" notions, with reference to the human mind (Sect. 9.9 and Fig. 9.9).

(i) Presenting the signification process, through which the mind attributes meaning to the signs in order to represent and transmit its thought content to other minds (Sect. 9.10); to do so, the mind must avail itself of a number of *Rings*, among which those:

- That control the correct form of a sign (Fig. 9.10)
- That identify a sign belonging to a signifier (Fig. 9.10)
- That interpret the meaning of an identified sign (Fig. 9.12)

(j) Undertaking a brief discussion of languages and the signs they are composed of, along with the 8 fundamental functions of languages (Sect. 9.11).

(k) Translating Bateson's model into the concepts of factual truth and falseness (Sect. 9.12).

(l) Trying to demonstrate how Bateson's model can be extended and operationalized to even take in the important topic of scientific explanations; scientific laws and theories as definitions and conjectural models of the world are produced by advanced mental processes (Sect. 9.13).

(m) Arriving at a definition of an intelligent cognitive system as one equipped with a mind that operates according to Bateson's hypotheses (Sect. 9.14).

The chapter ends with a brief summary of the explanation process and its operational closure (Sect. 9.15.2).

Chapter 10
Concluding Remarks: Toward a General Discipline of Control

An analysis of the history of technology shows that technological change is exponential, contrary to the common-sense "intuitive linear" view. So we won't experience 100 years of progress in the twenty-first century—it will be more like 20,000 years of progress (at today's rate). The "returns," such as chip speed and cost-effectiveness, also increase exponentially. There's even exponential growth in the rate of exponential growth. Within a few decades, machine intelligence will surpass human intelligence, leading to The Singularity—technological change so rapid and profound it represents a rupture in the fabric of human history. The implications include the merger of biological and nonbiological intelligence, immortal software-based humans, and ultra-high levels of intelligence that expand outward in the universe at the speed of light (Kurzweil 2001, online).

In Part I of this book I have tried, using the language of systems thinking to present the structure and typology of the control systems that allow man to create a livable world and to aspire to progress. Understanding the theory of control systems is necessary to attain a true control discipline, to introduce which I headed off in Part II on an ideal journey to guide the reader in learning to recognize and model the action of multiple *Rings* observed or even only imagined operating in various environments. This chapter concludes the journey into the logical world of control systems. At this point it is useful to present some final considerations to stimulate further reflection. In the first *section* below I shall attempt to outline several fundamental general *hypotheses* in order to propose a control discipline. These *hypotheses* lead to the conclusion that "*The world is conceivable as a Control System*" since its existence, as our own, is subordinated to a network of control systems, based on a typical multi-level holonic arrangement, that can be *viewed* as, and in fact *acts* as, a global control system. In concluding this work I also feel that it is necessary to touch on the *human aspects* of control. Control systems are useful instruments for man, who in most cases is the manager-user of such systems. We cannot ignore the human attitude to control; thus, I have left these considerations until the concluding *section*. A reflection on the content and limits of this study is presented as FAQs in the final pages of the chapter.

10.1   A Possible Control Systems Discipline

The exploratory path into the world of *Rings* undertaken in Part II (following the *systems thinking* point of view) allows me to suggest several generalizations, or conjectures, which, somewhat emphatically, serve to outline the structure of a hypothetical control systems discipline. In order to understand and justify these conjectures we must always keep in mind that the aims of a control systems discipline should be logical constructions—not physical apparatuses—which are not always immediately visible or understandable. A constant intellectual application is needed to master control systems, especially when these concern micro and macro environments too distant from human experience.

Let me start with the idea that if the world (as observed by systems thinking) is made up of repetitive and recursive systems of variables (Chap. 1) and actually exists and evolves in an orderly manner, except for pockets of disorder which man tends to eliminate, then it should be clear, even to a nonspecialist observer, that all the variables must be subject to some form of control. The variables of which the world is composed have limits or some objective value—determined by man or by the natural characteristics of the world—toward which they tend and to which they return when some disturbance moves them away from these values. A world without objectives, limits, or constraints that eliminate the disturbances to the component variables, both at the micro and macro levels, would be one dominated by chance, without any equilibrium or permanence. When disturbances move the dynamics of the variables away from the objectives, limits, or constraints, an adequate number of efficient *Rings* must exist to restore equilibrium, order, and stability and allow the world to exist in the form familiar to us all. Disorder and chaos are innate in many meteorological and geological phenomena, as well as in complex systems, but where possible man must control these dynamics and create order or at least protect himself from the consequences of disorder.

When this is not possible, it is not a question of the lack of *logical* control systems but of the difficulty in their practical realization through apparatuses sufficiently powerful and complex to deal with extreme dynamics. We need to only reflect on climate variables (Chap. 6) and on the disturbances to these from human activity in order to understand which wide-ranging *Rings* can regulate these equilibriums and how man tries to employ other control systems to eliminate or reduce these disturbances or seek refuge from the harm produced by the disturbed dynamics. If he is unsuccessful in these control actions and the climate variables attain other equilibriums, the world as we know it today could alter in an extremely damaging direction for mankind. New variables—heat, cold, sea levels, ozone and ultraviolet ray levels, etc.—must be controlled through systems that are increasingly vaster and interconnected. And what about the internal physiological controls and the external behavioral ones that account for our daily behavior?

From these premises I suggest the following conjectures, which go back in a systematic way to the concepts developed in the previous chapters:

First general conjecture: The world is composed of control systems.

According to the typical systems thinking perspective, the world is made up of variables, [Y], *from whose dynamics we can always identify or fix objectives or constraints,* [Y*], *whose achievement is made possible by the action of control variables,* [X]—*connected or unconnected—in balancing* loops. Ergo: *The world is made up of control systems interconnected in various ways.*

If we accept this first supposition and focus our attention on any type of control system, we realize that it is structured in a given way, since it belongs to a definite observational and operational context. Since the variables that make up the world are interconnected, by varying this context—zooming in or out—every control system is interfered with by other systems and the control must also extend to these interconnected systems. The control then becomes *multi-objective*; as a rule all multi-objective control systems must be strengthened by adding new levers. Thus, if we wish to increase crop yields and the birth rate in stock farming it is necessary to act on the availability of water by exploiting the naturally available quantities, thereby altering the flow of watercourses to create artificial reservoirs or drilling in underground strata. But this control system must take into account the drying up of surrounding areas. Thus the original control system acquires a new objective—not to dry up the surrounding areas—and must make use of other levers to provide the amount of water required by the arable land, for example, through aqueducts and desalinization plants.

Every *multi-objective* system is usually a *multi-lever* one as well. Each new lever that is added can, in turn, represent a variable to control for other connected systems; thus, the primitive control system is enlarged to take into account the necessary control on the levers that have been gradually added.

The *first conjecture*, together with these observations, allows me to derive a new conjecture:

Second conjecture: Control systems tend to become multi-lever and multi-objective.

The more we extend the boundaries of the reality to control, the more numerous and interdependent are the variables [Y]—*taking into account the system of objectives* [Y*], *the control levers* [X], *and the disturbance variables* [E]. *Ergo: The more we zoom in, the more the control systems become connected and interactive; normally, it is possible to set new objectives, add other levers, and identify new disturbance variables.*

At a given level of observation control systems are usually viewed as autonomous entities; however, we must accustom ourselves to thinking of them as interacting with other systems at the same level or at a superordinate or a subordinate level. The easiest means of interaction is the single-directional one, which occurs when the dynamics of *Y* in system A become a disturbance for system B or the bidirectional one when, simultaneously, A influences B and B influences A (interacting systems as discussed in Chaps. 2 and 4), with the emergence of *causal loops* of varying complexity.

Apart from the specific interactions, we can view control systems as classes of observed entities placed at different hierarchical levels. This can easily be seen when we consider the fact that many control systems, though independent, are

necessary for the functioning of others which, though independent as well, permit others at a higher level to function. In other words, there appears to be a holonic hierarchy (holarchy) among control systems (Sect. 3.8).

We can realize the different levels of control systems that exist as soon as we consider the fact that we can control our body's actions (movement, eating, etc.) because our organs function properly thanks to their control systems; and our organs, in turn, function correctly thanks to the cellular control systems that make up our tissue. It is equally clear that I can control my fingers for typing this sentence because I control my arm, shoulder, and entire skeletal muscle system, together with my sight organs; these words can flow from the keyboard because the neural system of the brain is under control; but this implies that the areas of the brain entrusted with the various cerebral functions are under control along with the individual neurons that have to control the impulses that arrive from "upstream" in order to pass on other impulses "downstream."

It is also clear that the control of the direction and speed of our car is possible because the various car systems and components, such as the gas level in the tank, turbo action, electrical system, and tires, are connected to specific control systems. Even the control system that tries to keep pollution emissions below certain thresholds implies, at the global level, the control of polluting companies, which limit individual emissions by controlling the energy consumption of their facilities and the maintenance state of their factories. How can infant deaths be controlled if we do not activate controls on health structures (obstetric and neonatal wards) and families (hygiene and diet)? How can we control health structures or families without controlling the level of preparation and education of the personnel? One of the problems in rich countries today is the control of Saturday night road accidents involving youth. It is clear that controlling the trend in such accidents implies controlling driving speed and the consumption of drinks and drugs. But to control speed and the alcohol level we need appropriate control systems, etc., which are part of a clear-cut hierarchy of controls.

In short, without taking anything away from the circular interconnections among control systems, it is clear that every control system at a certain level usually implies the existence of a lower level of control. These considerations lead me to the following conjecture:

Third conjecture: Control systems form a holarchy.

The world can be viewed as a holarchy of control systems at various levels. The systems at each level have their own properties, new or emerging; however, they are also influenced by those at a lower level while in turn influencing those at a higher level. Thus the world can exist because it is made up of a holarchy of control systems.

As mentioned in Sect. 3.8, the concept of *holarchy* comes from that of a *holon*, which was coined by Arthur Koestler (1967) to indicate any object or concept that can be observed from three points of view: (1) as an *autonomous* and *independent* entity that acts according to its own "canon" of behavior; (2) as a *superordinate* entity with respect to a certain number of component parts that it transcends and which displays emerging properties; and (3) as a *subordinate* entity that is part of a

vaster whole that conditions it (for a more in-depth treatment, see Mella 2009). Koestler considered in particular the class of holons formed by *vital systems* that display the dual tendency of self-survival and integration:

> Every holon has the dual tendency to preserve and assert its individuality as a quasi-autonomous whole; and to function as an integrated part of a larger whole. This polarity between the Self-Assertive and Integrative tendencies is inherent in the concept of hierarchic order; a universal characteristic of life (Koestler 1967, p. 343).

Holons can be arranged in a "natural way" according to interconnected *levels* that make up a hierarchical ordering, which can be either vertical (holons arranged *above* and *below*) or horizontal (holons arranged *before* or *after*). The vertical arrangement is called a holarchy and the horizontal one a holonic network. Koestler (1967, p. 344) defines *output hierarchies* as those that operate according to the *trigger-release principle*. In this case the *top holon* produces important processes that send signals to the base holons in order to condition their behavior. He defines *input hierarchies* as holarchies that operate based on the logic of successive filters; these produce progressive syntheses from the subordinate levels up to the superordinate ones, as if at each level the holons filtered or synthesized the inputs from the subordinate holons. It is clear that control system holarchies are simultaneously *input holarchies* and *output holarchies*, as shown in Sect. 3.8.

From an *ascending* observational approach—gradually zooming out—we can see that a miniscule pixel functions because a micro control system provides it with a certain state, since a superordinate Control system makes that state necessary in order to adjust the state of the monitor. But the state of the monitor derives, in turn, from the control objectives of the software that is operating at that moment, which in turn responds to the control needs of the operator who is using it, who in turn is influenced by the control needs of the group he or she operates in, which in turn is affected by the control system represented by the operational center where the group carries out its activities; this is part of a larger control system made up of the department, division, and entire organization, which in turn is controlled by other superordinate systems. The same applies for all organizations that can be considered control systems with regard to the achievement of the institutional goals for which they were created; the macro-level control directs (triggers) and depends on (filters) the control carried out by the organizational organs themselves, which direct (trigger) and depend on (filter) the control of the individual members, the processes, the operations, the movements, etc. This process involves a *descending* observational approach that zooms in from a broader to a narrower perspective.

We must now ask if observing control systems from a holonic perspective enables us to derive some generalizations. In my opinion the following supposition immediately arises.

Fourth conjecture: Control strategies and policies are necessary.

The higher the level we reach in the holarchy of control systems that make up the world, the more we encounter multi-lever, multi-objective, and interdependent control systems. The lower down we go in the holarchy, the more we find single-lever, single-objective, and independent control systems. Thus *every control system that is*

not single-levered and single-objective requires the definition and implementation of appropriate control strategies and policies.

In fact, the examples we have proposed show that the more control systems operate at a minute level—the base level (base holons)—the more their functioning is regulated by a single lever or by a limited number of levers. The further we move away from such base systems, the more they are structured with various control levers.

1. Holons, on successively higher levels of the hierarchy, show increasingly complex, more flexible, and less predictable patterns of activity, while on successive lower levels we find increasingly mechanized, stereotyped, and predictable patterns.
2. All skills, whether innate or acquired, tend with increasing practice to become automatized routines. This process can be described as the continual transformation of "mental" into "mechanical" activities (Arthur Koesler 1967, p. 346).

While a pixel, a radar gun, or a boiler in a condominium has only one objective and one control lever, the work of a computer operator, transport minister, or project manager involves controls that make use of multiple levers to achieve a number of objectives. We cannot control the monitor if we do not take into account the program and work environment. To reduce the number of highway accidents, speed controls are not enough; we must take into account the road system, the right to enjoy oneself, the needs of night clubs and of families, etc. We cannot control emissions only through controls on home boilers in a city; we must also consider other possible levers (traffic restrictions, replacement of obsolete facilities, congestion charges for cars, and so on). We also need to consider the cost and length of operation of the various levers, in addition to other objectives, such as minimum heating levels, commuter needs, and, lastly, social discontent.

The conclusion is clear: the higher up we go in the holarchy of control systems, the more necessary it is to specify a control *policy* as well as a coherent *strategy*. This conclusion, together with the *second conjecture of control*, implies that as the number of objectives to achieve increases so, too, does the number of variables to control; therefore, the levers that are employed must increase to allow for the control, and as a result the structure of the control system expands and becomes increasingly complex. What can we infer about the relation between the efficacy and structure of these control systems?

The fundamental principle of cybernetics gives us the following supposition:

Fifth conjecture: It is always possible to strengthen the power of a control system by adding new structural control levers. Structural levers are preferable to symptomatic ones.

Given a realistic objective, the control levers must be quantitatively matched *to the controlled variables and to the disturbances that condition the latter's behavior.* Thus *when the levers are not able to achieve the objective, given the disturbances, the control system can be improved either by changing the* capacity *of the existing levers (section and/or velocity) or by adding other control levers.*

The inadequacy of control systems in achieving their objectives does not always represent a structural defect in planning the number and capacity of the control levers; in many cases the system only apparently fails when unachievable objectives are set. In this case, rather than an overhaul of the system it is necessary to set more realistically attainable values for the objectives. The problem of the adequacy of the system with regard to the objectives clearly emerges when the objective, in addition to its quantitative value, also has a precise time reference that may not be achievable by means of the normal logical and technological knowledge available at that moment. It is useless to try to sink a ship by hitting it with a slingshot; the control lever (kinetic energy of the rock) is too small with respect to the variable to control (ship's mass); it is equally useless to hit a fly (mass of the target) with a cannon (energy of the projectile).

However, we must also consider that the choice of control lever depends on both the type of objective and the magnitude of the error that arises when the variable to control is far from the objective and the significance the manager attributes to this error. We must consider that not all possible control levers are also useful ones. Many systems become multi-lever when we add increasingly more efficient levers to substitute the *symptomatic* ones.

In conclusion, if the control system cannot take on the states of reality to be controlled through an adequate set of *structural* levers, then the control is destined to fail as soon as the real states can no longer be represented by the system. This leads to a further supposition.

Sixth conjecture: Control systems tend to improve their efficiency.

The holarchy of control systems is itself a control system that continuously produces new control systems that are more powerful and precise in their logical and technical structure, thereby improving the quantitative and qualitative performance. Ergo: Inefficient control systems inevitably tend to be replaced by more efficient ones through which new, more ambitious objectives can be achieved and stronger limits reached.

We know that progress in control systems derives in large part from improvements in the effector, detector, and regulator apparatuses that make up the physical structure of the systems. However, equally important is the contribution to the improvement of control systems from the continuous search for new logical structures of control that stem from the gradual improvement in our capacity to observe the world.

One obvious example of this is the strengthening and improvement in the control levers regarding infectious diseases from the discoveries of Edward Jenner, the English physician and scientist who, in 1796, was the first to understand the possibility of immunization against smallpox by means of the inoculation of a milder form of the smallpox vaccine (Riedel 2005). I do not know if Jenner was aware of the existence and understood the functioning of the powerful control systems of recognition-identification represented by the immune system (Sect. 3.2); in any case, control by means of the vaccine represented not only a strengthening of the preexisting immune system but, with respect to the knowledge available at that time, a new control system, more logical than biological, that was championed and

popularized by Louis Pasteur and systematically studied by Robert Koch, the pioneer of modern bacteriology, who, through his causal laws, enabled this important class of control system to be formalized. These laws are as follows:

1. The pathogenic agent must be found in all cases of disease.
2. It must be possible to isolate the pathogenic agent in the host and grow it in a culture, that is, in the laboratory.
3. When we transfer the pathogenic agent to a non-infected host we reproduce the same disease.
4. The same pathogenic agent must be isolatable in the new infected host.

And what about the new range of control systems that have been conceived of and developed after the discovery of Alexander Fleming, the biologist and pharmacologist who, in 1928, thought to attribute antibacterial capabilities to the Penicillium notatum mold, thereby deriving the famous penicillin which is one of the most powerful levers for controlling infectious diseases. Today the control of diseases in individuals has become common; nevertheless, man tends to set even loftier objectives, aspiring to a global control of infectious diseases while moving toward the objective of "zero disease" over the entire continents. In fact, cholera, plague, leprosy, and malaria have almost disappeared from our world, being confined to areas not yet reached by any control system.

Let us now consider another important chapter in the strengthening of control systems: inventions and progress in mechanics and computer science and their incorporation into robotic control systems. We have come very far indeed from Leonardo da Vinci's "lion robot," of which there is no physical trace, though there are references in his manuscripts together with suggestions for realizing his project for a robot soldier! According to historians (Rosheim 2006), this lion robot was commissioned from Leonardo to celebrate the entry into Lyon of the French king, Francis I. Perhaps in homage to this city, Leonardo designed, and according to historians constructed, a semi-movable lion programmed to walk toward the king, stop in front of him, and rise up on its hind legs while opening its chest with its front paws to display a lily (symbol of France) in place of its heart (Atlantic Codex, f. 812; Madrid Codex I f. 90 v).

And who can ignore the qualitative leap in mechanical control systems due to the invention of the chip that permits digital control. Who could have imagined the progress that has taken man from the initial woodturners to modern robotic turners with laser control? Finally, who could have imagined the power and improvement in control systems that have permitted space probes to leave the solar system or the Hubble telescope (which will be replaced by the James Webb Space Telescope) to identify the most distant galaxies?

These considerations lead us to a new *conjecture*.

Seventh conjecture: The holarchy of control systems produces a continuous improvement in controlling our world.

The holarchy of control systems that make up our world produces a continuous improvement, allowing us to increase the variety and strength of the possible

controls and the scope of the attainable objectives. Ergo: *From a systems thinking perspective, the perpetual improvement in control systems permits the control of every micro and macro variable that makes up our world.*

We can view the holarchy of control systems as a control system that produces progress. There is no global desire to produce global progress; continuous improvement is not the result of decisions by some supreme authority but derives from the invisible hand of the unknown action of control systems that improve other control systems, which in turn improve other control systems as part of a circular causal chain that in the end produces the global improvement we are experiencing.

Echoing Koestler, we can thus consider the holarchy of control systems as an *open hierarchic system*, a *machine* made up of a holarchy of holons interacting at different levels which can survive thanks to the action of control systems.

> I have tried to explain in it the general principles of a theory of Open Hierarchic Systems (O. H. S.), as an alternative to current orthodox theories. It is essentially an attempt to bring together and shape into a unified framework three existing schools of thought—none of them new. They can be represented by three symbols: the tree, the candle and the helmsman. The tree symbolizes hierarchic order. The flame of a candle, which constantly exchanges its material, and yet preserves its stable pattern, is the simplest example of an 'open' system. The helmsman represents cybernetic control. Add to these the two faces of Janus, representing the dichotomy of partness and wholeness, and the mathematical sign of the infinite ... and you have a picture-strip version of O. H. S. theory (Koestler 1967, pp. 220–221).

This system of control systems produces general progress in life through the two-dimensional improvement—upward and downward—in the holons-control systems, as if there were a *Ghost in the Machine*, thereby producing an inevitable evolutionary process of improvement (Banathy 2000).

This leads to the final supposition:

Eighth conjecture: The world is conceivable as a control system.

The holarchy of control systems inevitably tends to transform itself into a network of multidimensional, multi-level control systems that act as a resistance to disorder and disturbances. Ergo: *The world in its unity and totality can be viewed as a single control system.*

Control systems tend to be holonically interconnected in an holonic network of relations that, by favoring local control, produce global resilience of the network as a whole. The network functions if all its Rings (in the nodes) act, according to their appropriate times and in an appropriate space, simultaneously and in a coordinated way, revealing a selfish behavior aimed at their survival within the network. The network can withstand damaging events and a lack of resources, replace local control systems with other ones, and repair the damages caused by external disturbances in the form of natural calamities. The network tries to repair the damage or restore the functionality of the destroyed control system by replacing those parts that do not improve efficiency. In other words, the network tries to survive (Arthur et al., 1997; Barabási 2002).

More inefficient control systems cease their activities, and others are created which produce a higher level of control, extending control to multiple objectives.

Political and legislative superstructures favor the creation of new and efficient systems that improve and expand the holonic network. In other words, the network tries to survive by eliminating every form of disturbance that can endanger the homeostasis of our autopoietic variables.

> We define survivability as the capability of a system to fulfill its mission, in a timely manner, in the presence of attacks, failures, or accidents. We use the term system in the broadest possible sense, including networks and large-scale systems of systems (Ellison et al. 1997, 97).

The holonic network of control systems is, from a theoretical point of view, capable of achieving higher objectives (improvement), as part of an unending interaction among individual lower level control systems, which are connected hierarchically and in a reticular manner to systems at vaster levels. It is precisely due to this global structure of the network of control systems that the local interactions among systems are increasingly intense and robust and produce global homeostasis acting as an "emergent algorithm."

> An emergent algorithm is any computation that achieves formally or stochastically predicable global effects, by communicating directly with only a bounded number of immediate neighbors and without the use of central control or global visibility (Fisher and Lipson 1999, p. 6).

> An emergent algorithm may produce results that are distributed globally but do not exist locally, that are entirely local, or that are represented both locally and globally (*ibidem*, p. 4)

Increasingly concrete and evident is the risk that the local control systems which, at various levels, constitute this network, will be managed in a conflictual manner, thus leading to the achievement of inopportune objectives. There is also the risk of a dearth of managers capable of dealing with the challenges posed by the complexity of structural interactions.

I shall say no more, since any further comment would no longer be technical in nature but political.

10.2   The Human Aspects of Control

As we know from Sect. 3.1, in many circumstances man not only represents an element of the control systems (artificial or natural) that surround him but also normally interacts with the *control processes* those systems carry out. In artificial automatic systems the interaction is minimal: it is manifested by a simple positioning of the objectives, with man playing the role of *governor* of the system. The maximum interaction occurs when man is an integral part of the entire *chain of control* and, using his limbs, senses, and decision-making capacity, replaces the effector, detector, controller, and communication systems.

Moreover, as a *detector*, man recognizes the *errors* in the vital variables with respect to the internal homeostatic constraints, attributing to these *errors* the meaning of unpleasant *sensations*: pain, thirst, hunger, cold, etc., which lead him to make urgent

control decisions, often aimed at eliminating the symptom (unpleasant sensation), rather than to intervene on the appropriate variable to control (Sect. 5.6). Even more direct is human intervention in many social systems and in organizations, as will be illustrated in Chaps. 7 and 8.

There are many aspects that characterize the *interaction* between man, as manager-governor, and control systems:

1. Fallibility
2. Learning
3. Inventive and innovative capacities
4. The search for active control
5. Intolerance toward passive control

Fallibility is the first and most evident characteristic of man's interaction with control systems. Understood as the human trait of making mistakes and committing errors, thereby compromising the control system's operation and the success of control processes, fallibility takes on different expressions and nuances. It can derive from an *inability* to understand how to regulate the control levers, recognize the correct variables to control, or measure the error, as occurs with children when they do not know how to move their effectors in order to use a pencil to draw a shape or with people who are driving a car for the first time and do not know how to regulate the steering wheel to guide the car or park it. Fallibility due to "ignorance" is overcome with practice and thus with *learning* and the gaining of experience. Fallibility can also be the result of various *impediments*, such as disease, disabilities, and distraction, which limit man's possibility of action regarding the apparatuses that form the system's chain of control. "I could not hear the alarm because I was listening to the stereo"; "My eyes were tearing up, so I couldn't see I was about to hit the wall"; "I had drunk too much and my reflexes were too slow"; "I was on the telephone and couldn't turn off the tap in the bathroom, so the house was flooded"; "Too tired, too distracted, too fast, too little light, too tiring, too changeable, etc." are all examples of factors that hinder efficient control.

I will mention other examples. The Greek general (manager) orders the standard bearer, in the heat of battle, to display the flags (which are codified) that give the cavalry the order to attack. However, because of the noise the standard bearer, who is in charge of communications, mistakenly orders the infantry to retreat from battle, thereby causing the army's defeat. Spotting an enemy submarine approaching, the commander orders the helmsman to "steer to starboard," but the latter has been injured and carries out the maneuver too slowly, thus causing the ship to be hit by an enemy torpedo. The head oarsman of a trireme (regulator), who calls out the rhythm to the rowers, receives the order from the commandant (manager) to increase the pace of the strokes in order to avoid a collision (detection of distance and speed) with the enemy vessel; however, the rowers (effectors) are exhausted and cannot keep up with the pace, and thus the ship is rammed. The motorcyclist (manager) encounters a very tight curve but, due to his or her speed, is not able to regulate the motorcycle's inclination (control lever); thus, he does not succeed in (effector) negotiating the curve, ending up off the road.

These few simple examples clearly show that human fallibility can intervene in an evident or subtle manner in all phases of the control system to produce an inefficient functioning of all the apparatuses in the chain of control, including the system for the transmission of information. There are three basic lines of action to mitigate the fallibility of man's intervention in control systems.

The *first* is obviously the proper *instruction* of the manager-governor regarding the features of the various elements in the control system. The preceding *sections* of this chapter indicated the control system features that can represent a source of failure; we must know what these are and avoid them. The *second* line of action is to provide the manager-governor with the necessary support to facilitate his or her interventions: typically, the *error signals* (e.g., acoustic alarms for distance, rear video cameras for cars, speed limit controls, detectors of moving obstacles), *calculation systems* for calculating the optimal regulation, automatism in the use of the effector, etc. and the other *algedonic signals* that warn about the impossibility of achieving the objectives. The *third* and most drastic form of reduction in human fallibility is the complete substitution of man in control systems by transforming the latter into automatic systems (errors by surgeons are avoided by robot surgeons; driving errors are avoided by automatic systems; braking errors are avoided by automatic braking systems; pilot errors are avoided by an automatic pilot; etc.).

To compensate for his *fallibility*, man possesses a *learning capacity*; he is able to *learn*, that is, to *control* himself by transforming himself into a true control system of his own actions in order to eliminate errors committed as the *agent* of the control. *Learning* by the manager is manifested, in successive cycles of a control system, as the *ability to become more and more effective in eliminating the error* through more precise and timely regulation decisions. Learning increases the manager's *experience* and improves his or her capacity to detect, regulate, and implement the operations needed to control the dynamics of the passive variables, as shown in the model in Fig. 10.1, which refers to the general model in Fig. 3.1, to which has been added the control systems of the managers themselves.

Whenever control is carried out by man, the ability to correctly maneuver the control levers is essential for an increasingly more precise and effective control. When man is part of the *effector* apparatus, learning is in the form of *training*, mental and muscular improvement, to acquire *automatism* in actions. When man is part of the *detection* apparatus he must prevent or mitigate the decay in the error perception organs by correcting focusing problems through proper lenses, maintaining the hearing and olfactory apparatuses, and so on. When man intervenes as a *regulator*, he must avoid alterations in judgment, a decline in the capacity to calculate, by resisting the taking of drugs that can inhibit these capacities or by using drugs that can block the decline in the cognitive functions needed for regulation. There is no need to provide abstract examples; everyday life testifies to the importance of learning and improvement in controlling our driving, handwriting, gesturing, postures, and relationships with our neighbors.

To facilitate learning and consolidate the influence of experience in human control, the most effective line of action is the standardization of the means of man's intervention through the formalization of procedures, strategies, and, more
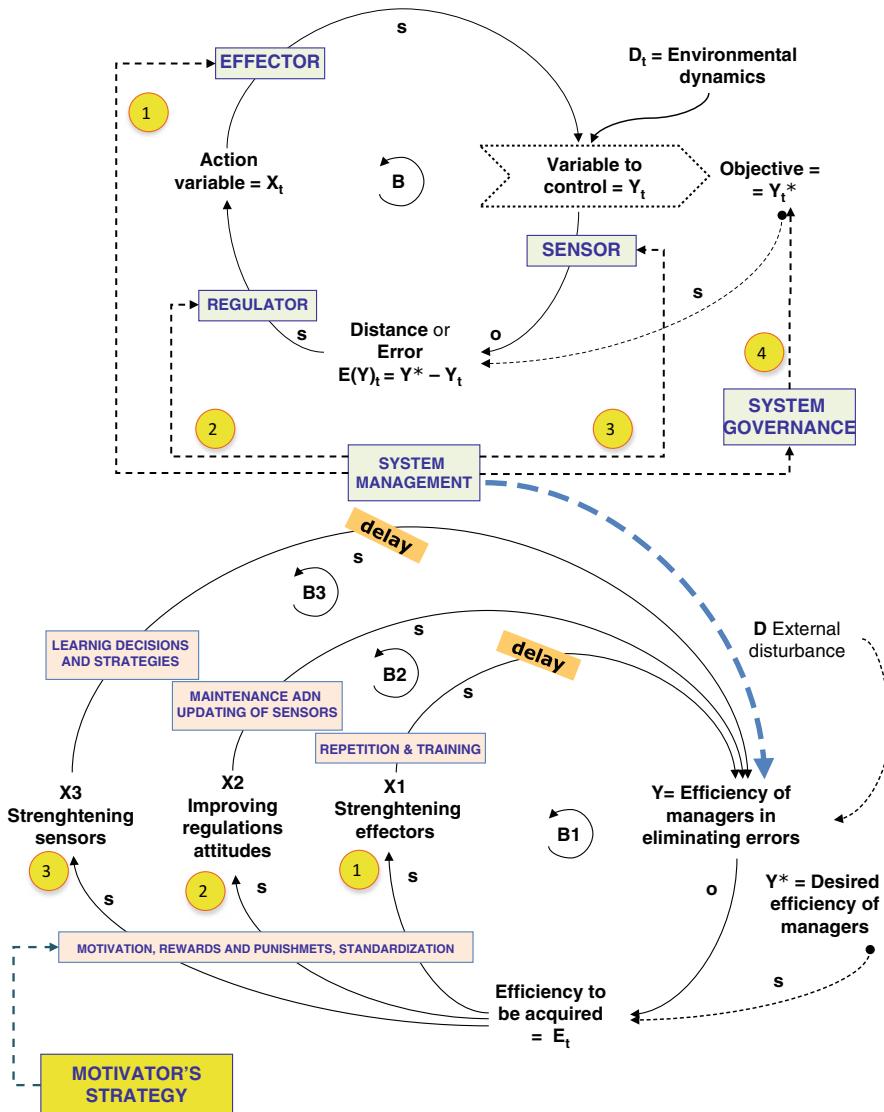
Fig. 10.1 Managerial learning for the efficiency of control (starting from Fig. 3.1)

generally, *routines*. Section 4.5 introduced the concept of *routine* (or plan of action) to indicate a consolidated strategy the manager feels is efficient to apply to the manual control of a control system that involves numerous cycles or to regulate similar control systems in all similar control contexts.

The consolidated and formalized *routines* facilitate learning on the part of various individuals involved in managing similar control systems with regard to the variables to control [*Y*] and the levers to use [*X*]. In learning and applying the

routines, even the manual control tends to become similar to an automatic control for the manager-governor, since the routine indicates to the manager how to regulate the lever without having to calculate and decide anew, on each occasion, the variations in [*X*] and the maneuvers of the apparatuses that form the chain of control.

I can still remember how difficult it was, during my first driving lessons, to guide that formidable system of multi-objective and multi-lever controls in the car. Regulating the wheel to turn right, but not before regulating the directional indicator, and at the same time regulating the speed in adequate proportion to the driving maneuver, regulating braking by judging how much room there is to stop. How many uncertainties and repetitive actions there were! Gradually, by following the *routines* recommended by the driving instructor and learning how to apply them, and at times improve them, I learned to drive a car, and now, after years of driving, it does not seem possible that I had to learn to apply so many *routines*. Subsequently I decided to get a pilot's license for a lightweight airplane. The control of even a small plane requires the application of a myriad of routines. There is a routine for every regulation, every lever, every objective, and every external disturbance. Check the runway and wind, listen to the instructions from the control tower, take off at the right speed, check to see the sky is clear, go at full throttle, gain speed, pull back on the control stick, rise to 450 feet, fold in the landing gear, give feedback to the control tower, rise to the proper altitude, avoid the off-limit flying zones, regulate the engine revs, check the direction of flight with the compass and correct your route, correct the lateral movement if there is lateral wind, check the temperature of the cylinders, check the sky up ahead, gain altitude, descend if you are too high up, etc. For each control lever there are smaller routines that indicate how many levers to use at the same time, which maneuvers have priority, etc.; all these controls represent the normal flight routine, which has to be repeated for each flight, and most of which must be repeated at regular intervals, many times, for each flight.

Which routines do we follow each morning to get ready to go to work? Which routines do chefs follow to prepare the various dishes ordered by customers at the same table? There is probably no need for other examples to understand that routines are the result of the *learning* needed to reduce the *fallibility* that characterizes man's interaction with control systems.

Learning is typical in most animals; many of their controls are guided by innate routines; for example, all spiders know which levers to activate and in which order to individually weave a spider web; all squirrels know the routine to collectively follow in building a small dam. Many routines in the animal kingdom produce unstoppable control automatisms. The lowly dung beetle pushes balls of excrement following an innate routine that obliges it to use its back legs by walking backward; it continually repeats this innate routine by always moving its legs in the same way, simulating the pushing movement even when the excrement ball is taken away from it (Russell and Norvig, p. 39).

Normally there is no learning in mechanical or physiological automatic control systems. On the contrary, in such systems there is usually a decline in performance over time. To combat this there is need for *maintenance* of the efficiency of all the

apparatuses that make up the chain of control or at least for *prevention* to slow down the gradual decline in the system's operation. Nevertheless, learning can also be found in artificial automatic systems, especially in recognition control systems (Sect. 3.2), where the regulation is carried out through neural networks that "learn" how to maneuver the control levers based on the frequency and type of error, reacting increasingly more efficiently. A typical example is the modern video game, where a human player competes against the routines carried out by a computer which, with its "artificial intelligence," learns to react with increasing precision, thereby forcing the player to also learn by devising game strategies, that is, his or her own personal routines.

A third characteristic distinguishes human intervention in control: *conceptual* and *innovative* behavior, in other words, man's capacity to identify and conceive of new control levers to strengthen systems, thereby further increasing his experience and laying the groundwork for new forms of learning, as part of a reinforcing *loop* that is continually acting to produce the *progress* we can all observe in the world.

There is no comparison between the rudimentary control levers of primitive sail boats with rudder bars and the modern servomechanisms of the giant cruise ships that carry 5,000 passengers and punctually control room service, meals, entertainment, etc. Think about the leap in quality in the control of picture taking. The first cameras had completely manual controls, with imprecise, difficult-to-regulate levers. Today's modern reflex cameras have made man the *governor* who decides when and where to take photos. The *manager* has become automatic and regulates focusing, shutter speed, shutter openings, and other levers with the utmost precision, thereby always guaranteeing quality photos. Consider also the vastly improved quality in modern plasma and LCD television technology compared to the first sets with manual controls; and how far we have come from wood fireplaces and stoves to the modern thermoconvectors that regulate the desired temperature for each room in our house separately. It is easy to recognize the improvements in the levers for the technological, social, biological, and medical controls that make our lives easier in an increasingly crowded and complex world. All these improvements are the result of man's capacity to continually invent more sophisticated control systems and of the continual innovation in the apparatuses that allow these systems to operate.

After this rapid analysis of the interactions between man and control systems, several considerations about the emotional aspects of human intervention in *control processes* are in order. Many control systems in which man is the manager-governor are designed to control the behavior of other men, in the broadest sense of the term. We can thus distinguish between *controller* (or active) *subjects* and *controlled* (or passive) ones. *Controller subjects* are individuals, groups, and organizations that represent the *active elements* in the control processes; as governors they set the objectives of the control and design and produce the control systems in order to control the behavior of the other *controlled subjects*, according to a predetermined policy. *Passive subjects* of control are those that must be controlled in, among other things, their actions, behavior, and interactions while respecting the constraints in the achievement of their objectives.

Applying the first rule of systems thinking and zooming in or out, it becomes clear that man, when observed as a *controller subject*, wants to extend his control (at times, his *dominion*) to cover every controllable variable regarding the behavior of other men but also, and more generally, regarding his own body, environment, the animal, and vegetal world as well as the physical environment in the broadest sense. Through his conceptual and learning capacities man can continually identify new control systems and increasingly more powerful control levers in order to travel to deep space and control the weather, desertification, global warming, disease, social groups, minutest parts of his cells, and even biological clocks that determine his longevity.

I propose the following *first supposition*: Thanks to his intelligence, there appears to be a "natural" (even innate) tendency in man to control every variable that impacts his existence as an individual, a member of a group, and as a species. We can paraphrase John von Neumann's quote at the beginning of Chap. 2 and state that if "*All unstable processes we shall control*," it is also true that "man *can*, and *must* control all the unstable processes in his world," particularly those that can benefit his existence and well-being.

Through science and technology man can conceive of, design, and produce evermore powerful and precise chains of control of all kinds. Aided by mathematics, quantum physics, space, nuclear, genetic and computer technology, nanotechnologies, materials science, and all other forms of invention and discovery, man can design multi-lever and multi-objective systems which are increasingly vast, complex, and flexible in order to control not only himself and his own environment but also the infinitely small and infinitely large elements in his world.

If we consider man as a *passive subject* of control, the postulate in the first supposition appears to become inverted: man displays a natural aversion to being controlled. Rules are followed only if forced on him, and, in any event, they are not accepted willingly. Man is clearly "rebellious" toward all forms of coercion, and it is also evident that he wants to be the *controller* of the behavioral variables, not *controlled* by them. This attitude comes out strongly in revolutions or in wars against an enemy that wants to subjugate a population; but it can also be affirmed by all those that unwillingly accept undergoing periodic medical checkups, submit tax returns, or obey the rules set by organizations and social groups. And it is the attitude even of the most observant religious individuals who must follow the precepts imposed by their church.

In light of these observations, we can posit the following *second supposition*: Man displays a "natural" *resistance* to being a passive subject of control processes. Thus, the aforementioned quote by John von Neumann could be completed as follows: "man *wants* to control all the unstable processes in his world, but he possesses a natural *resistance* to being the passive subject of the control processes that concern his own behavioral sphere."

A final observation of the human aspects of control regards *self-control*. I have presented numerous examples of *self-regulation* regarding variables related to man as a psychophysical entity: the control of equilibrium, temperature, needs, aspirations, handwriting, etc. However, now I wish to deal with *self-control*, a term I use

to indicate the control exerted by man over his behavior, lifestyle, moral and ethical choices, etc. as well as the capacity to achieve given standards of behavior or directly desired objectives chosen from other possible objectives. Self-control is not an innate attitude; man experiences instincts and reacts to these often in an unconscious manner. Self-control modifies human behavior to produce desired actions and reactions, which are often imposed on man. The defensive instinct leads most of us to turn away from the sight of blood and to avoid touching snakes, mice, even insects. And yet millions of persons have activated forms of self-control to modify their conscious behavior: doctors, nurses, and first-aid personnel treat wounds every day; biologists, researchers, and even children will themselves do not fear contact with certain animals.

Many of us instinctively react physically and verbally to violence and oppression. A good deal of self-control is needed to forego such instinctive reactions and rely instead on legal remedies. Many people are content with their present state, while others try to improve their situations even if this entails great sacrifice. Not to mention the self-control of practitioners of Yoga and meditation, who regulate the rhythms of their breathing, move their limbs into incredible positions, and even control their states of consciousness. In conclusion, self-control enables man to consciously modify his state and the quality of his behavior for a host of reasons, as the following quote clearly illustrates:

> Self-control refers to the capacity for altering one's own responses, especially to bring them into line with standards such as ideals, values, morals, and social expectations, and to support the pursuit of long-term goals. Many writers use the terms self-control and self-regulation interchangeably, but those who make a distinction typically consider self-control to be the deliberate, conscious, effortful subset of self-regulation. [. . .] Self-control enables a person to restrain or override one response, thereby making a different response possible. (Baumeister et al. 2007, p. 351).

Clearly self-control follows the general model of every control process. It is necessary to define the behavioral variables to control, Y, and the objectives, Y^*, as well as to specify the apparatuses that form the chain of control. Figure 10.2 proposes a simple model of a system that carries out the self-control on a single behavioral variable. There can be multiple and varied control levers. The model in Fig. 10.2 indicates three such levers: the resources, commitment, and effort acceptable as levers through which the individual can implement the self-control. Each lever could be individually specified and have different effectors that provide different actions for improvement with different action functions, g. When the self-assessment detects a gap between the desired quality of behavior and the actual behavior, thus producing *insatisfaction*, the individual must then vary the state of the levers through a reaction function h, which represents the level of *motivation* and the *cultural state* of the individual; this reaction function can modify his or her *willingness* to provide greater resources, commitment, and effort to improving his or her state.

It is not only athletes, scientists, researchers, religious personalities, or ascetics that produce continuous forms of self-control; anyone possessing self-esteem and the esteem of others will wind up reflecting on the quality of his or her behavior,
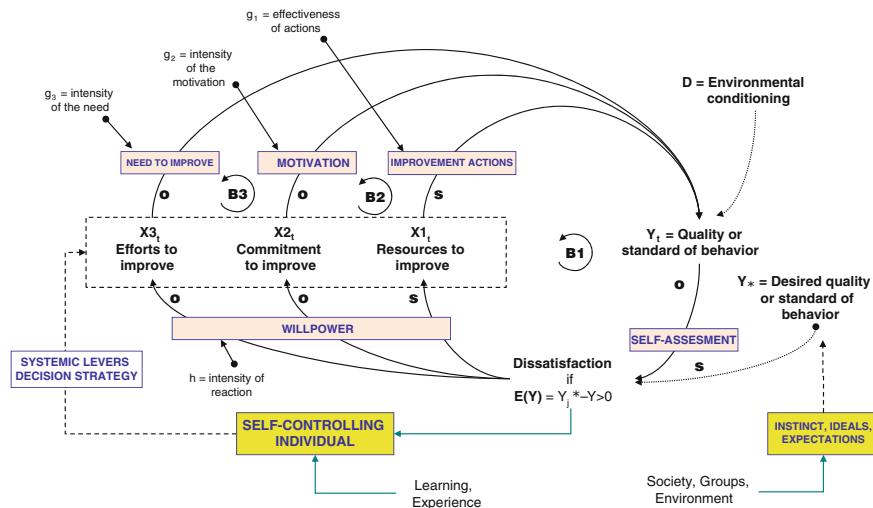
Fig. 10.2 The general model of self-control

attitude, and relations and undertake self-control, even though other alternatives are possible which often are more convenient and easily realizable, though less stringent and imperative.

As a conclusion to the above observations, I would suggest completing the first two hypotheses on man's attitude toward control with the following auspicious ending: "if man *wants* to control all the unstable processes in his world but possesses a natural *resistance* to being the passive subject of the control processes that concern his own behavioral sphere, he should at least be capable of controlling himself."

10.3   Conclusion: Frequently Asked Questions (FAQ)

After this long journey I have recommended the reader undertake to help him or her explore the magic world of control systems, it is useful, as an aid to further reflection, to go back and present several basic themes as recurring questions that anyone might ask himself or herself or me.

Are the control systems we have studied in this book different from those presented in books on the theory of regulation and control in engineering?

Absolutely not, even if they seem different, since I have employed a different language and different diagrams to represent them. As shown in Sect. 2.16.10, engineering books generally use rectangular boxes which contain the processes, or functions, that transform the inputs, X, into the outputs, Y, which are indicated on the arrows. Moreover, in engineering books on control systems, the objective, Y^*, is

usually considered as the fundamental informational input of some "technical machine" that must achieve the objective, to which the *deviation* is added via the feedback line. Our model, on the other hand, emphasizes the relation between the control levers, X, and the controlled variables, Y; the objective, Y^*, enters into the calculation of the deviation and is represented on the right, usually next to Y. Finally, engineering books deal with real systems, mechanical, electrical, hydraulic, etc., that is, they deal with *technical* control systems made up of actual apparatuses. This book instead presents control systems as *logical models* which must be constructed in the real world by a "control chain" of technical apparatuses. I have considered some concrete control systems—the radio, shower, bathtub, car, submarine, etc.—in order to indicate models which are easy to understand and remember.

Are they then different from those presented in mathematics books on systems theory?

Here, too, the answer is no. In this book I have presented the *logic* of control without recourse to formal mathematics, preferring simple linear systems (when possible) with constant coefficients or systems similar to these. System theory texts—or those on system science—deal with abstract systems which are usually based on the theory of differential equations that connect X and Y. Math books that examine systems also consider a third variable, the *internal state of the system*—whose variations are intermediate to those of the inputs and outputs—by studying the relations $[X \rightarrow \text{internal state}] \rightarrow [\text{internal state} \rightarrow Y]$. The internal states of systems are not explicitly considered in this book since the relations between X and Y have been directly described by means of the simplest *transfer functions*.

The text presents models of very simple real systems. Are these typical models, that is, prototypes?

The answer is yes. The aim of this book is not to explain control systems which are designed or constructed in concrete situations occurring in the physical, biological, and social worlds, but to propose several simple models to aid the reader in observing and understanding the magical functioning, the wonderful complexity, and the incredible variety of control systems that enable our world to exist and progress. How do control systems function, and what is their significance? These are the questions I hope the material in this book will provide answers to. Precisely for this reason I have chosen to utilize quite elementary, at times trivial, models of control systems, but ones which everyone knows and has experienced. Each of these elementary systems has a unique characteristic that influences the control process. Thus, they can be considered prototypes of elementary though general systems. The *control of the level of liquids*—in the example I used, a cistern—is the prototype of control systems with dependent and constrained levers. The *control of velocity*—in our example, referring to a car—represents the prototype of multi-lever and multi-level control systems that require a control strategy. Finally, the control of *three-dimensional movements*—in our case, a traditional airplane and a Cartesian robot—can be the prototype of multi-objective systems that require a control policy. The numerous variants, together with the other examples (including those that concern the interconnections among systems), can always be linked to the preceding prototypes.

Why is the book divided into two parts?

Part I, entitled DISCOVERING THE RING, presents the logic of control systems and proposes a minimal typology. Part II, THE MAGIC OF THE RING, tries to illustrate the significance of the ubiquitous presence of control systems in all observational contexts and at every possible level we might zoom in on. Borrowing Charles Morris' (1938) three dimensions to describe languages (syntactic, semantic, and pragmatic), I suggest interpreting Part I as the *syntax* of control systems and Part II as the *semantics* of these systems. The *pragmatics* of control systems, understood as the answer to the question "How do we do things with control systems?" can be perceived by intuition throughout the book or can be recognized by observing the real world. In fact, the *Rings* are instruments for achieving an objective, respecting a constraint or staying within a limit. They exist and are constructed to allow for the ordinary operations of any kind of process.

Can the control systems presented explain everything?

The answer is yes, but under the condition that the reader succeeds in applying the models to both the macro and micro environments presented in Part II, following the general rule of systems thinking that recognizes the importance of zooming out and in to observe and understand the world. Though elementary in nature, the prototypes I have proposed are useful models for understanding the logic of control as well as for identifying the means by which this logic can be translated into concrete observations, more so than into concrete realizations. They are not operational models and cannot explain everything, but they do offer a general and broad vision of the problem of control from a systems thinking point of view.

The text also presents simulation programs. Are these in any way operational?

The simulation programs are above all illustrative and not immediately operational. The models are elementary just like the control systems they simulate; they do not take into account the numerous complications and disturbances that can present themselves in the real world. Their aim is simply to numerically verify the coherence of the models presented. For "small" systems I have preferred to build the models using Excel—which I think everyone is familiar with, at least its elementary commands, which were the only ones adopted—and Powersim, one of the many available simulators for "large" systems. I admit that the reader might prefer setting up my simple programs in a different manner to make them more elegant and complete. Nevertheless, the more expert reader could easily simulate the control systems by building *ad hoc* programs using any personal program language available today (from the immortal C to C++ to Java). Even MATHEMATICA, *"The world's ultimate application for computations,"* as written in the Wolfram Research Web site (http://www.wolfram.com/mathematica/), allows us to simulate highly elaborate control systems by using differential equations and differences at any level of complexity.

What is innovative about this book?

The book presents several innovative features. Formally, the most evident is that it sets the control process within the framework of the general theory of systems, which is presented following the systems thinking approach. In fact, the book borrows its terminology and models from systems thinking. In terms of substance, the book seeks to launch a message more than to present a technique.

What is the message of the book?

Even though we are not accustomed to seeing them, control systems dominate our world, our life, and our very existence. "Seeing" control systems is not difficult, but we are not trained to think in terms of control. The control process takes shape in one or more balancing *loops* that are produced by control systems of every type, in every context, and for every purpose. To observe them we must accustom ourselves to systems thinking; but, as many have pointed out, teaching people to think in terms of *loops* is not widespread in our educational system, which is instead oriented—in terms of both science and techniques—toward teaching how to identify chains of open causal relations, along the lines of Kantian thought. The *message is therefore* if you truly want to understand the world, use systems thinking and, above all, think in terms of control systems. Search for the *loops* as well as the objectives, constraints, or limits toward which the variables tend. Finally, try to understand which variables are the control levers and how they operate to reduce the distance from the objectives.

Is control theory a science or a discipline?

Picking up from the previous message, here is another piece of advice: Do not be content only to observe control systems; you must get used to *thinking* in terms of *control processes*. To do this, you must train yourself, apply the concepts you have learned on a daily basis, and broaden your observational experience. In short, you must consider the observation of the world—in terms of control systems—as a *discipline of control* that will help you improve your intelligence and expand and improve your capacity to create models to understand and to explain and forecast phenomena. The simpler the basic instruments are, the more we will understand the control mechanisms; the more we understand the control systems, the more we will apply ourselves to the control discipline; the more advanced we are in the discipline, the more desire we feel to learn sophisticated instruments of understanding and simulation from books on engineering, math, physics, biology, economics, sociology, etc.

Is this why the text favors simpler control systems?

Yes. You cannot undertake the discipline of gliding by trying to pilot a glider without instructions and proper training. Similarly, unless you are an engineer or a mathematician, it is not possible—or, if possible, certainly not profitable—to attempt to understand the control methodology in terms of differential equations and complex feedback. I have preferred an elementary, intuitive, visual, gradual, non-mathematical, and logical approach to control, using the language of systems thinking, The reader who studies the control discipline "will know" when it is the right moment to move on to more sophisticated techniques.

Are there other topics the book does not deal with that should be studied?

The answer is certainly "yes." In Part II of the book I have tried to present a large variety of control systems that surround us and appear in the most varied environments: within ourselves as individuals, in our homes, in our towns, in animal populations, in society, and in organizations. I have even quickly delved into the world of weather phenomena. My aim was to explore control systems that allow

man as a psychophysical being to exist, briefly examining the control systems that could serve to operationalize Bateson's model of the mind. Though I have tried to examine various topics, many have had to be omitted due to the difficulty in entering into technical details as well as the vastness of the control problems that I would have had to deal with. For example, I have not considered the physiological control systems of the human and animal body in general and its organs in particular. Any physiology book on humans and animals is so voluminous as to make it difficult to choose the topics to treat. Neither have I taken up, even in part, those control systems activated by the endocrine system; apart from some brief comments in the preface, I have also avoided examining control systems that regulate the functioning of our nervous system and brain to produce thought and behavior. The knowledge regarding chemistry, electrochemistry, and neuronal and endocrinal physiology is so specialized that translating this into control system models would not have been very effective. The control systems chosen for presentation in this book therefore do not exhaust the types of systems or the possible observational contexts. There is much more also to be written on social and organizational control systems as well as on physical and biological ones. Finally, having chosen a didactic, non-mathematical approach, several important aspects have been neglected, three in particular: the problem of the stability of several multi-lever control systems which are interconnected (interfering showers, for example); the behavior of nonlinear control systems; and the relation between control theory and problem solving. Though I have simply noted this last point, the topic deserves more detailed treatment.

The text does not mention the problem of controlling the future, that is, the process through which the present produces the future but the future conditions the present. Can we control the future?

This is a perplexing question. Perhaps I have not explicitly written that there is a mutual relationship between the present and the future, but was there any need? The entire control discipline assumes this relationship; in fact, it is based on this. Let us reflect a moment: The idea itself of *objective* indicates our capacity and will to orient our actions toward future states we desire to achieve; the idea itself of a control lever assumes the possibility of acting "today" to produce the dynamics that help achieve "tomorrow's" objectives. Asking if we can control the future makes no sense in the context of control theory; by definition, control always involves the future (objectives), which conditions the present (control levers), which produces the future (movement of Y toward Y^*). In any event, in many parts of the book I have explicitly considered the processes of *change* and *decision-making* (and thus of the desirable or the necessary future) as control systems. Therefore, the control concept is intrinsically linked to the concept of future.

With extreme synthesis, what is the essence of the control discipline?

As I have observed several times in the book, control systems are at the basis of the order, and thus the existence and progress, of the world, which is understood here as the system of interconnected and interacting variables that must be

subjected to some control system. Very succinctly, the discipline of control proposes three main teachings.

First: If we truly want to understand the world it is necessary but not sufficient to observe it in terms of systems of interconnected and interacting variables (systems thinking); we are also obliged to recognize and understand the operational logic of the control systems that condition the movements of those variables toward the achievement of some objective, constraint, or limit.

Second: Control systems occur at multiple levels and in ever-broader contexts; the higher up we go in the hierarchy of control systems, the more powerful these become. Starting from single-lever and single-objective systems—typical of the more minute biological automatic control systems—we see that these progress, at higher levels, to artificial multi-lever and multi-objective systems which are increasingly more powerful and flexible. For this reason, we can think of control systems as elements of a typical dual-directional holarchic arrangement.

Third: There is a close relation between problem solving—which concerns the malfunctioning of the system—and control systems, since, in general, problems derive from the lack of control systems, the inadequacy of the existing ones, or their malfunctioning. When we do not correctly apply control systems to the variables in our life, then we are the cause of our problems. If we encounter problems we must always ask the following: Which control system did not function and why? Which control systems can we apply? How can we improve the control process?

For these reasons, the discipline of control must accompany us all through life in order to accustom us to identifying the objectives (constraints, limits), Y^*, and the dynamics in the Y variables that make up our life, and to help us understand which levers, X—guided by the amount of error $E = Y^* - Y$—can determine the dynamics of any system, taking into account the external disturbances. Only in this way can we control our future.

10.4   Complementary Material

10.4.1   Control System in Religions: Approaches in Catholicism and Buddhism

It is not easy to find a shared definition of "religion." However, taking a cue from ancient Western tradition, Cicero's definition appears acceptable to me: "Religion is all that concerns the care and veneration directed at a superior being whose nature we define as divine."[1]

---

[1] "Religio est, quae superioris naturae, quam divinam vocant, curam caerimoniamque effert" (Cicerone. De inventione. II,161).

In general, all religions can be distinguished on the basis of three variables (the three "C"): the "creed," that is, the salvific end religion is thought to allow us to achieve; the "cult," that is, the practices adopted to carry out the salvific behavior that guarantees we will achieve the goal of salvation; and the "Church," that is, the set of structures (places of worship) and individuals (the clergy: priests, gurus, imams, etc.) that in various ways control the diffusion of the "cult," which allows the faithful to achieve their "creed" of salvation. I hold that every religion is structured in such a way that the "believers" must and are able to control their behavior so that it is in keeping with the optimal salvific behavior specified by the "creed" and controlled through the "cult" supplied by the "Church." This logic also applies to all kinds of "salvific practices," even those of a nonreligious nature, as, for example, in Buddhist practice.

Every religion and salvific practice thus includes (or in itself can be considered as) a multi-lever control system whose general model includes a *structural* control and also at times a *symptomatic* one. In such a control system the objective is a "minimal (sufficient) state of perfection" and the variable to control is the "actual state of existence"; the gap between these two states represents the "imperfection" to be eliminated (defined qualitatively and often also quantitatively), and this can be interpreted and denominated in several ways in the differing religions: sin, sorrow, disobedience, blindness, offense to God, etc.

When the individual is aware of being in a "present state of imperfection" he or she must try to eliminate the "imperfection" (error) by means of control levers established (or suggested) by the "cult" proposed by his "Church." These control levers are normally modifications in the quality (and often in the quantity) of the present behavior in order to bring the "present state of existence" closer to the "minimum state of perfection." We can distinguish between levers that act in direction "**o**," or *anti-evil levers*, in that these try to eliminate the negative characteristics of behavior that inhibit the "present state of existence" from reaching the "state of perfection" set as the objective; and levers that act in sense "**s**," or *levers of good* (broadly speaking and to be specified on each occasion), which improve the "present state of existence" not only by eliminating the imperfection but also by bringing the present state to levels that are superior to the objective, thereby producing a "state of holiness" (grace, illumination, inner peace, etc.).

Figure 10.3 describes the general model I have outlined. The system is normally multi-lever, since it relies on a multiplicity of *structural* levers of the X1 type (in favor of good) and the X2 type (against evil). It is likely that once the error (imperfection) is eliminated the manager-believer will continue to activate the "levers of good" by giving $E(Y)$ a negative value. In this case the error takes on the meaning of "perfection," and the "state of perfection" produces some form of "holiness," whose meaning is specified differently in each religion.

Many religions have "rules of the Church" that permit recourse to *symptomatic levers* in order to *quickly* eliminate the error without activating the structural levers "for good" and "against evil" (which are difficult to activate, as their action is very *slow*). For example, the Catholic religion has the sacrament of Penitence (remorse, confession, and good intentions), which leads the believer to temporarily eliminate
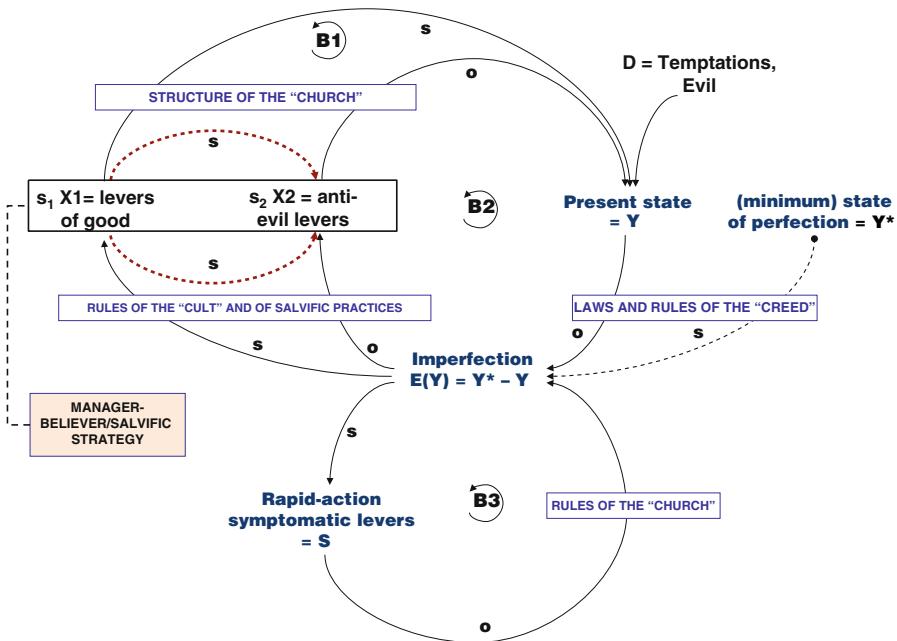
Fig. 10.3 General model of a control system of a religion

his or her state of sin (imperfection). In Hinduism bathing in the sacred rivers in the sacred cities is permitted (on special occasions), the Ganges being the most famous, as a way of obtaining forgiveness for one's sins and as an aid in achieving salvation. Similarly, going on a pilgrimage to Ayodhya, an ancient city on the right bank of the Gaghra River, and bathing in the river can annul even the worst acts imaginable (for example, killing a Brahmin).

Without pretending to be complete, but only to suggest a guide for the reader, I will try to apply the model in Fig. 10.3 to outline the control system inherent in the Catholic religion and in Buddhist practice.

In my opinion, the Catholic religion is based on four principles well known to the believer: (1) sin exists (as does evil); (2) one must be "healed" from sin and maintain oneself in a state of grace in order to merit eternal life in God's presence; (3) there are ways to redeem oneself from sin and acquire and maintain a state of grace; (4) the Catholic church represents the "vehicle" (teachings, practices, and structures) for eliminating sin from the life of the faithful and for allowing him or her to gain eternal life. A few brief considerations are in order here (I apologize for the summary and partial nature of what follows).

The first principle should be entirely evident: Sin is an offense to God. We are born with original sin, and Jesus has taken on Himself the sins of all mankind, redeeming the world through his death on the cross, thereby reconciling man to His Father: "*God the Father, 'sending his own Son in the likeness of sinful flesh and for sin, he condemned sin in the flesh*" (Encyclical Letter "Dominum et Vivificantem"

of the Supreme Pontiff JEAN-PAUL II, 1986, Catechism, online). Sin appears to be unavoidable:

> Man cannot avoid having at least venial sins, as long as he is alive. Nevertheless, he must not give too much weight to these sins, which are defined as venial. They seem insignificant when you commit them, but how frightening when we add them up! Many little things put together form weighty things: many drops fill up a river, just as many grains of sand a heap. What hope remains then? First and foremost, Confession should be made ... (Sanctus Augustinus, In epistulam Iohannis ad Parthos tractatus, 1, 6: PL 35, 198).[2]

The second principle recognizes that the ultimate end of the believer in the Catholic religion is to attain and maintain a state of grace in order to merit eternal life in the presence of God, who sent Jesus Christ, his only son, to take upon Himself the sins of the world for the salvation of mankind and to die upon the cross: "*He will come again to judge the living and the dead*" (The Apostles Creed); "*He will come again in glory to judge the living and the dead, and his kingdom will have no end*" (The Nicene Creed): "*Immortality" in a Christian sense means the survival of man after his terrestrial death, for the purpose of eternal reward or punishment*" (Encyclical Letter "Mit Brennender Sorge" of the Supreme Pontiff Pius XI, 1937, Catechism, online).

> But to live always, without end—this, all things considered, can only be monotonous and ultimately unbearable. This is precisely the point made, for example, by Saint Ambrose, one of the Church Fathers, in the funeral discourse for his deceased brother Satyrus: "Death was not part of nature; it became part of nature. God did not decree death from the beginning; he prescribed it as a remedy. Human life, because of sin ... began to experience the burden of wretchedness in unremitting labor and unbearable sorrow. There had to be a limit to its evils; death had to restore what life had forfeited. Without the assistance of grace, immortality is more of a burden than a blessing". A little earlier, Ambrose had said: "Death is, then, no cause for mourning, for it is the cause of mankind's salvation" (Encyclical Letter "Spe Salvi" of the Supreme Pontiff Benedict XVI, 2007, Catechism, online).

The third principle is that there is a remedy for sin; there are types of behavior to *avoid* because they lead to sin and behavior to *pursue* because it leads to the annulment of sin and to the state of grace. The fourth principle states that Jesus, through his Church, teaches people the "paths" to take to avoid sin or to be "cured" from sin in order to acquire and maintain a state of grace.

It is not easy to list these "paths," but generally speaking they concern behavior to be avoided or improved in order to avoid sin, along with behavior to undertake and maintain to achieve a state of grace (Catechism, online). Above all, it is necessary not to disobey the Ten Commandments (loop [**B1**] in Fig. 10.4) and to avoid the six sins against the Holy Ghost and the four sins that call out for revenge in the presence of God (FrancoBampi, online). Moreover, it is necessary to avoid

---

[2] "Non potest homo, quamdiu carnem portat, nisi habere vel levia peccata. Sed ista levia quae dicimus, noli contemnere. Si contemnis quando appendis; expavesce, quando numeras. Levia multa faciunt unum grande: multae guttae implent flumen; multa grana faciunt massam. Et quae spes est? Ante omnia, Confessio ...."
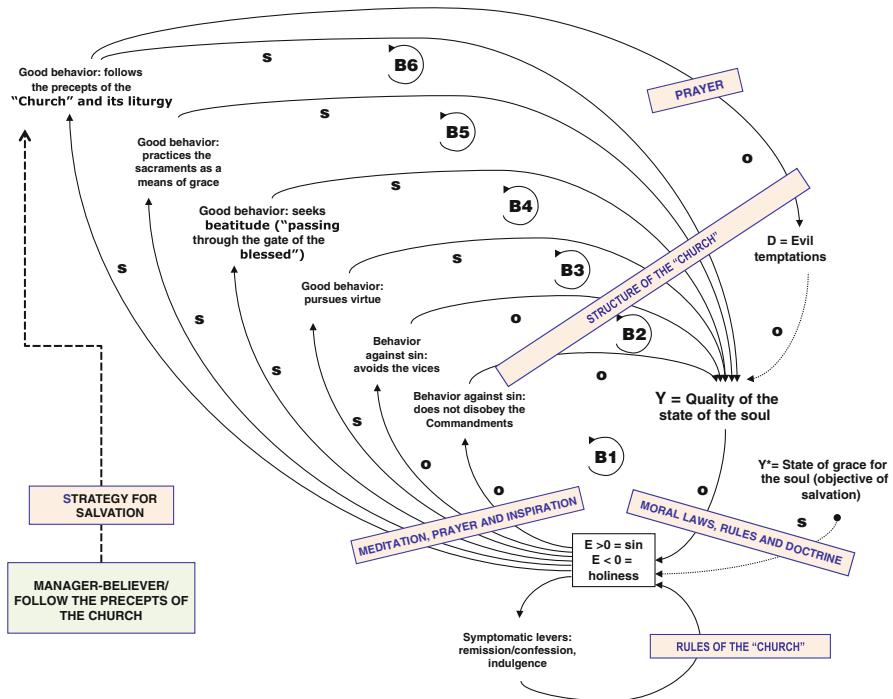
Fig. 10.4 The pluri-lever control system of (in) the Catholic religion (a simplified model)

the seven deadly sins: (1) pride, (2) greed, (3) lust, (4) wrath, (5) gluttony, (6) envy, and (7) sloth (loop [**B2**]). There is a multitude of types of virtuous behavior (state of grace); an initial list is indicated in the formation of loops [**B3**] and [**B4**]. Loops [**B5**] and [**B6**] elevate the state of the soul through the typical practices of the Church: the sacraments, precepts, and liturgy.

Figure 10.4 only shows the main loops, since many other ways of gaining salvation can be found in the teachings of the Church and in the behavior of the saints. It may happen that a believer not only has no sins and avoids behavior that could lead to sin but also behaves in such a virtuous manner as to lead his or her soul to a state above and beyond what is required for eternal salvation. In such cases the error is in fact negative and denotes a state of holiness.

The model of the control system in Fig. 10.3 can be applied not only to religions but also to the "salvific practices" that did not originally represent a religion in the strictest sense. I shall try to apply it to Buddhism in Fig. 10.5, asking indulgence of the reader for the summary nature of the treatment. The very heart and core of Buddhism represent the logical schema of the "Four Noble Truths," put forth by Buddha in 588 B.C., which represent the essence of the teaching (Dhamma or Dharma). They are considered as "truths" because they are *real* and form an *incontrovertible* fact of life. We may be aware or unaware of them, but they really exist in the world; Buddha simply revealed them to mankind (see The Dalai Lama, 2010 online).
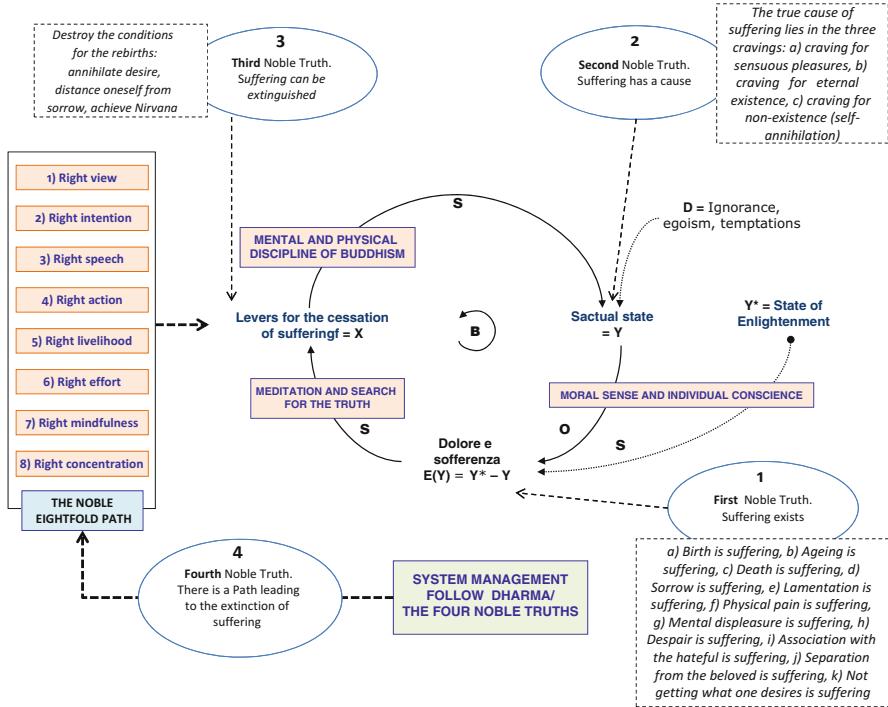
Fig. 10.5 The pluri-lever control system (of) in Buddhism (a simplified model)

In order to better understand the significance of the "Four Noble Truths" we must keep in mind that Buddhism did not arise as a religion but as a life practice within Hinduism. The "Four Noble Truths" are (The Dalai Lama 1998) as follows:

1. *Suffering exists*; this is a universal truth: (a) Birth is suffering; (b) ageing is suffering; (c) death is suffering; (d) sorrow is suffering; (e) lamentation is suffering; (f) physical pain is suffering; (g) mental displeasure is suffering; (h) despair is suffering; (i) association with the hateful is suffering; (j) separation from the beloved is suffering; (k) not getting what one desires is suffering.
2. *Suffering has a cause;* the origin of suffering is craving, which passes on from birth to birth; accompanied by pleasure and greed, it finds increasingly renewed delight in the here and now. Craving is of three kinds, namely, craving for sensual pleasure, craving for eternal existence, and craving for nonexistence.
3. *Suffering can be extinguished* by cutting off its causes through the complete fading away and extinction of its origin, craving. The extinction of suffering is obtained through the complete elimination of desire, the separation from longings. The state in which the complete cessation, or extinction, of craving, ignorance, suffering, and the cycle of involuntary rebirths (saṃsāra) is achieved is Nirvana (or Nibbana). Once Nirvana is attained the chain is broken: life and death no longer have any meaning once the state of enlightenment is reached. The Buddhist conceives of Nirvana not only as the liberation of the soul from

rebirth after death but also as the state of perfect peace and happiness and of imperturbable serenity and the absence of sorrow. For this reason, achieving Nirvana is for the believer the ultimate aim in life, the supreme state of enlightenment in which liberation from sorrow is achieved and the cycle of reincarnations is interrupted;

4. *The path leading to the extinction of suffering* is the "Noble Eightfold Path" composed of eight steps or factors: (1) right view, (2) right intention, (3) right speech, (4) right action, (5) right livelihood, (6) right effort, (7) right mindfulness, and (8) right concentration.

Considered from the standpoint of practical training, the eight path factors divide into three groups: (1) the moral discipline group, made up of right speech, right action, and right livelihood; (2) the concentration group, made up of right effort, right mindfulness, and right concentration; and (3) the wisdom group, made up of right view and right intention. These three groups represent three stages of training: the training in the higher moral discipline, the training in the higher consciousness, and the training in the higher wisdom (Bodhi 1999, online).

10.4.2   Control of Chaos and of Chaotic Dynamics

Figure 7.7 (Chap. 7) presented five populations forming an ecosystem whose chaotic dynamics became regular and periodical as a result of an artificial control that set *maximum* or *minimum* constraints for the number of individuals. These constraints can be respected through selective hunting to reduce the number of individuals in excess of the maximum or through planned repopulation to increase the number of individuals when they fall below the minimum. Figure 7.7 allows us to reflect on the possibility of extending control systems to chaotic dynamics in general. I would suggest that chaos is a widespread phenomenon, at least in theory; however, we usually are not aware of this since we tend to eliminate or reduce its effects with appropriate controls. I shall first clarify what notion of chaos I am referring to.

If we set aside the mythological or the cosmogonical notion of chaos—as the disordered primordial universe from which our known ordered universe was formed—then the term chaos was first formally used by the American mathematicians, James Yorke and Tien Lien Li, in a 1975 work entitled Period 3 Implies Chaos, which dealt with differential equations, in particular the logistics function (Kellert 1993). We can thus limit ourselves to the notion commonly used in the mathematical theory of systems, according to which deterministic chaos is the characteristic of a system's dynamics that, although regulated by relatively simple recursive functions, such as a quadratic equation of the type

$$x_{n+1} = cx_n(1 - x_n), \tag{10.1}$$

has a temporal dynamics apparently devoid of cyclicity, or regular occurrence, which is sensitive to the initial conditions, so that modest variations in these conditions can produce a temporal dynamics that is totally different and unpredictable.

C=	3.920
x(0)=	0.889

MEAN	0.5242
MAX x(n)	0.9784
MIN x(n)	0.0827

C=	3.925
x(0)=	0.891

MEAN	0.5255
MAX x(n)	0.9785
MIN x(n)	0.0827
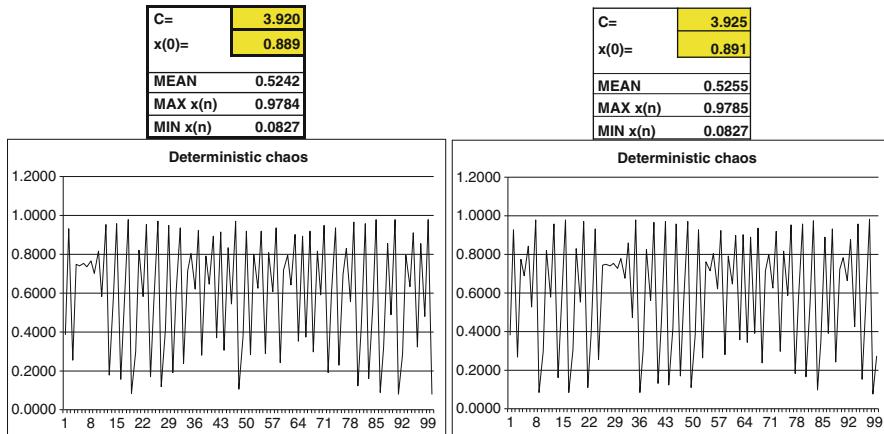
Fig. 10.6 The chaotic dynamics of a quadratic function with two different initial conditions

This definition of chaos as a characteristic of the dynamics of a recursive system describes the so-called deterministic chaos (Flake 2001; Gleick 2008). A similar definition can be used for multi-agent systems where individual behavior is dependent on a change of state individual probability function. The dynamics of this system—derived from the synthesis of individual behaviors—is unpredictable, without particular cyclicity or recurrences. In this system the chaos is defined as probabilistic, and it is revealed because the individual behaviors are regulated by specific probability functions, one of which is the so-called tent map (Mella, 2007). Figure 7.7 shows a chaotic system in population dynamics, since the simulation model immediately demonstrates that modest variations in the initial population conditions (initial numbers, rates of increase and decrease, even for a single population) produce entirely different dynamics.

If we refer to Fig. 10.6, which indicates the chaotic dynamics of the function (10.1), and assign certain initial values to "x_0" and "c," we can clearly see the dependence of the dynamics on the initial conditions. When the conditions $c = 3.920$ and $x_0 = 0.889$ change slightly, becoming $c = 3.925$ and $x_0 = 0.891$, the dynamics are completely different. Therefore, for certain initial values the quadratic function is chaotic, since it is sensitive to these conditions and does not display any particular cycles (at least for the period considered in the simulation).

The meteorologist, Edward Lorenz, became aware of this feature of chaotic dynamics almost by chance when, simulating the dynamics of a quadratic function using various pairs of initial values, he observed that for some pairs the dynamics were in fact chaotic, thereby concluding that meteorological processes, being sensitive to initial conditions (pressure, humidity, temperature, etc.), can have chaotic dynamics for particular values of the initial state parameters.

Figure 10.6 also shows the maximum and minimum values of the chaotic function for the period of reference. It is natural to then test how the dynamics of the function would vary if we introduced a simple control: for example, setting the
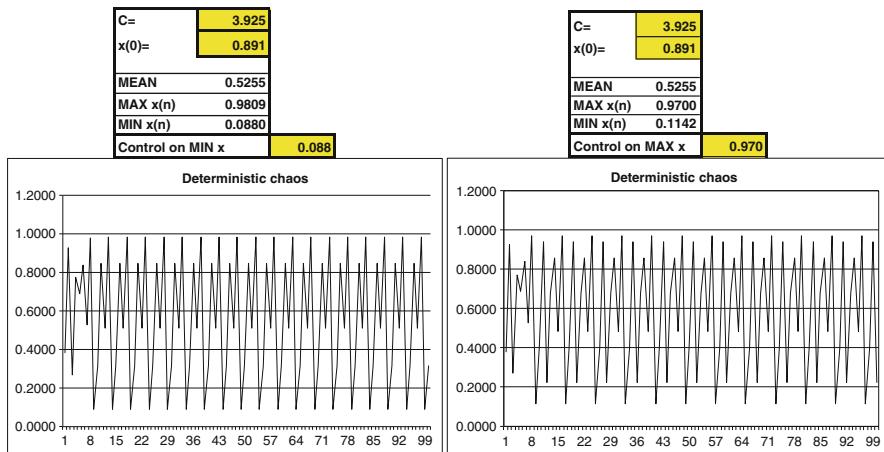
Fig. 10.7 The non-chaotic dynamics of the quadratic function with a maximum or a minimum control

condition that the function cannot exceed the *maximum* value $x_{\max} = 0.950$ or, alternatively, the *minimum* value $x_{\min} = 0.088$. With these very limited maximum and minimum controls (for simplicity's sake applied one at a time) the dynamics soon become cyclical and the chaos vanishes, as Fig. 10.7 clearly shows.

It is precisely the ease with which we can keep chaos under control, even with loose constraints, that allows us to understand why chaos has almost disappeared from our daily lives, remaining only where the controls are not applied (for more, see Flake, 2001: Chap. 13). Thus, it would be easy to eliminate chaos in the whirlpool of a stream by simply building a bank along the sharpest curves. The chaos in the behavior of pedestrians that run into each other on a crowded sidewalk (for a simulation, see Legion, online) could be kept under control by appropriately indicating separate paths for both directions using signs. The chaos of a crowded bus stop, the entrance to crowded locales, or the departure gates at airports could be eliminated by a ticket system indicating the order for those who want to gain access.

10.4.3   Variant Structure Control Systems

The control systems presented in the previous chapters were always of the *invariant structure* or *invariant action* kind: the operating logic behind both the number of levers as well as the technical apparatuses (Fig. 2.10) remained unchanged during the control process. In other words, the *structural characteristics* of the "technical system" (Sect. 2.9) were kept unchanged; these characteristics are summarized in the *action* function, $g(Y/X)$, and the *reaction* function, $h(X/Y)$. These functions were assumed to be unchanged during the control cycle. To make the hypothesis of invariance more evident, these functions were in fact given the most elementary
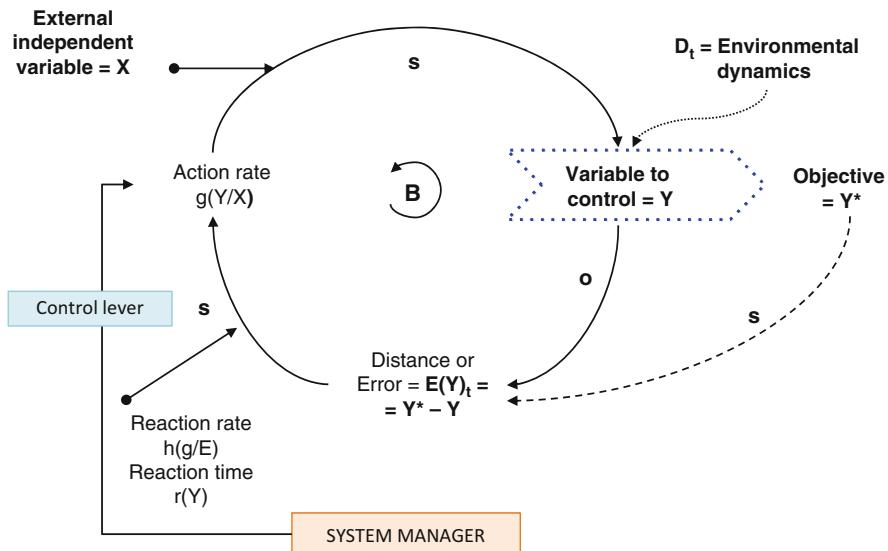
Fig. 10.8 An elementary variant action control system (reformulation of Fig. 2.2)

form, that of the "action rates" and "reaction rates," "*g*" and "*h*," respectively. In other words, they were constants independent of X, Y, and t. Nevertheless, in Sect. 5.4 we saw that the structure of a control system can be modified through the internal or the external strengthening of the system (Sect. 5.4, Fig. 5.2).

These structural modifications make control possible even when the manager cannot directly act on the variable X to modify the value of Y and eliminate the error, $E(Y) = Y^* - Y$, but must necessarily modify the action and reaction rates. Thus, we have a different form of control in which, since X cannot be manipulated, the manager must try to eliminate $E(Y) = Y^* - Y_t$ by acting only on the rates "*g*" and "*h*" (more generally, on the structure of the action and reaction functions), so that, even with the values of X unvaried, the desired values of Y_t are consistent with Y^*. We can refer to such systems as *variant action* (or *structure*) control systems, due precisely to the fact that the values of Y depend not on variations of X but on values of the rates "*g*" and "*h*," which are varied to allow the varying processes (black boxes) to carry out the control in the presence of a flow of X that cannot be modified. The general model of a single-lever invariant action control system, presented in Chap. 2, Fig. 2.2, can be reformulated as shown in Fig. 10.8, where the action rate "*g*" itself becomes the *control lever* to be modified (assuming that the rate "*h*" has been readjusted); X is no longer the control lever but instead represents an *external variable* such as D.

Figure 10.8 can be reformulated in the more significant and understandable manner shown in Fig. 10.9, which, while appearing quite different than Fig. 10.8, is in fact entirely equivalent. The advantage of transforming Fig. 10.8 into Fig. 10.9, by isolating the variable X, allows us to see more clearly that X is a variable
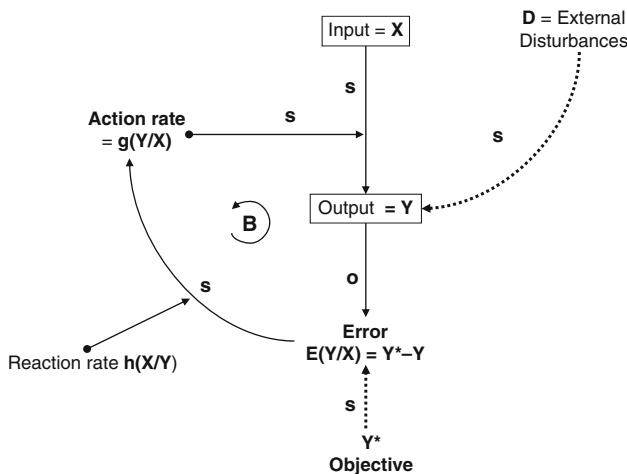
Fig. 10.9 General model of a single-lever variant action control system (reformulation of Fig. 10.8)

considered as an independent "input" of the system and no longer as a control lever. In order to ensure that the input X will produce a Y that tends toward Y^*, the control system must act to suitably vary the action rate $g(Y/X)$. The advantage of Fig. 10.9 is even clearer when we have to construct dual-lever control systems whenever Y depends on two or more external variables, $X1$ and $X2$, acting in parallel based on their respective action rates, g_1 and g_2. The model in Fig. 10.9 would then expand to become the model in Fig. 10.10.

The model in Fig. 10.10 represents the variant structure of a typical dual-lever control system whose levers can act jointly (Sect. 4.1) or separately (Sect. 4.2). Thus, in systems where the two levers act "in parallel" there are three possibilities for controlling Y (that is, for eliminating the error): acting on g_1, on g_2, or on both. The choice obviously depends on the type of situation the model represents; in any event, a precise control *strategy* must be defined.

We can formulate a third standard model which assumes that the systems operate "in series," or "sequentially," as indicated in Fig. 10.11, which clearly shows that we are again presenting a dual-lever control system since, during the initial phase, the input X is transformed into the output Y1 according to a variant action rate, g_1. Subsequently, Y1 becomes the input which the system transforms into the final output Y2, according to the specific variant action rate $g_2(Y_2/Y_1)$ for that phase. The error is still calculated with reference to Y2, but its elimination entails the threefold possibility of acting on g_1, g_2, or both; as always, the order of the reaction defines a control *strategy*. As I have already observed, in the models in the figures above, the reaction rate "h" could actually be modified to allow the control.

By observing the models in the figures above, we clearly see that, beyond the exchange of the variables that represent the control levers (X in the models in
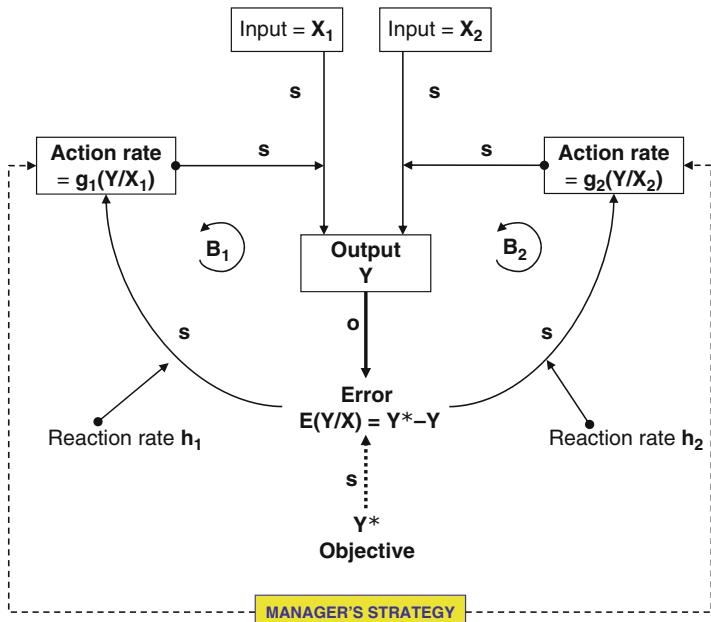
Fig. 10.10   General model of a "parallel" dual-lever variant action control system

the previous chapters, the rate "g" in the models in this *section*), the control logic is the same: acting on the *control levers* to move Y toward the objective, thereby eliminating the error. The fundamental problem which the system "designer" must face is determining the action rate "g" needed to eliminate the error without being able to intervene on the quantity of flow of the X variables. The solution to this problem can be found relatively easily for the basic systems in the models in the above figures; however, for more complicated structures, such as systems of *composite* transformation with more than two inputs and numerous links, "in series" and "parallel" at the same time and over various layers, the regulation of the values of "g" can turn out to be quite difficult.

I hold that variant action control systems constitute perhaps the simplest models for representing a wide class of physical or logical systems, among which:

– Transformation systems
– Production systems
– Neural networks

Though we cannot give due consideration here to this topic, some brief observations are nevertheless in order. I define *transformation systems* as all those apparatuses that transform an input variable, X, into an output variable, Y, with the "convention" that Y represents the result of some *qualitative* transformation (pieces of meat vs. minced meat; dirty dishes vs. clean dishes, etc.) or *quantitative* transformation (a lamina with thickness "a" vs. a lamina with thickness "b";
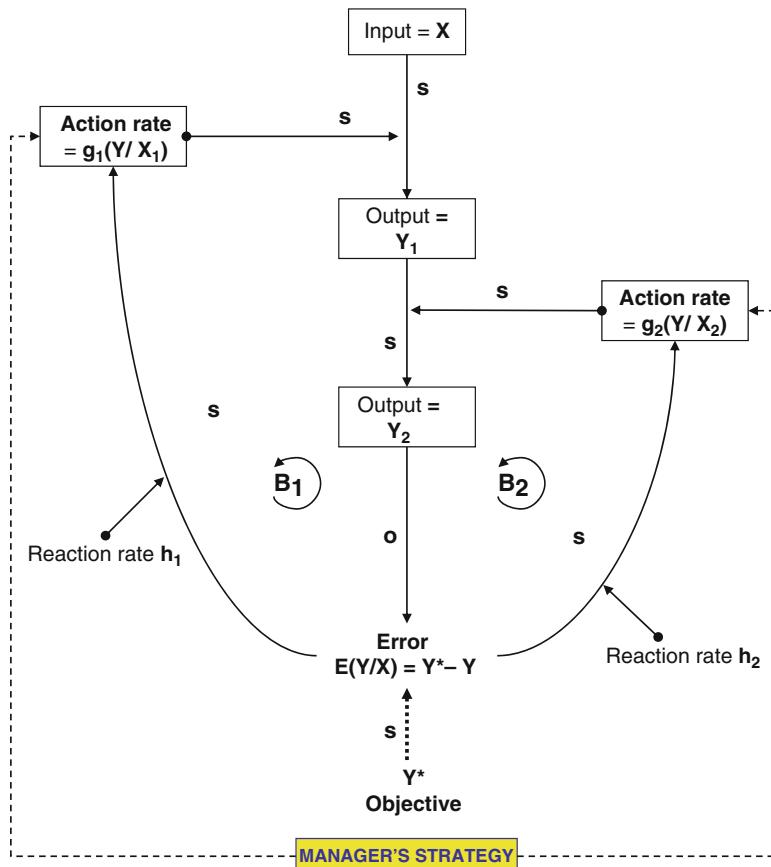
Fig. 10.11 General model of an "in series" dual-lever variant action control system

investment of an amount of capital "A" and disinvestment of capital in the amount of "$A + I$", etc.) of X itself, based on action rate "g." In many cases the control system cannot act on X, the value (or variable) that must be transformed, which represents an external input, but instead needs to modify the action rate "g" so that the transformation produces the desired Y. The rate "g" takes on the significance of "transformation coefficients."

Production systems can be viewed as a particular class of transformation system. Through appropriate production apparatuses they produce the various combinations of factors of production flows, X, in order to obtain flows of finished products, Y, which can be considered in all respects as the result of the transformation of factors according to the productivity rate "g." When the control system cannot achieve the production objectives by maintaining constant the productivity rates and modifying the factor amounts (control levers) X, which arrive at irregular intervals (for example, consider sick people in a hospital), the productivity rates must be modified by *varying* the productive structure.

Neural networks can also be viewed as particular transformation systems—which act through modules which are linked and structured over various layers—which, through successive processes ("in series" and "parallel") modify the values of the independent inputs, X, into values of output, Y, by carrying out operations on X or on the sublevels of Y through appropriate coefficients of transformation known as "weights" (which normally are between 0 and 1). These weights can be fully equated with the action rate "g." If the neural network has an objective Y^* and, through the inputs X and the "propagation" process, the network produces a value Y which differs from Y^*, then the transformation coefficients (weights) must be modified, since it is not possible to act on the inputs to the input modules in order to gradually move Y toward Y^*. In this sense the neural network learns to "recognize" Y^* given the entry values X, which cannot be directly controlled. Therefore, even the neural networks, no matter how they are designed, can in principle be conceived of as variant structure control systems; in this case they can learn through the feedback processes, which modify the weights of the links among the upstream and downstream nodes. Today the neural networks normally do not have a control system for the weights, which instead are determined randomly based on reiterated "backpropagation" processes.

Transformation systems are usually *composite* since their structure has more than two inputs arranged over various layers. In order to implement the structural control for *composite* transformation systems it is necessary, as always, to define a strategy, among which we can mention the following:

(a) *Simple strategies*, which act on a single lever represented by a chosen transformation coefficient; the aim is to try and identify a single "g" such that its modification immediately eliminates $E(Y)$. This strategy is not always possible since there may not exist a g_i which, by itself, and taking into account the range of admissible variations, succeeds in bringing Y toward Y^*. The easiest strategy is to act on the final layer—let us say, M—considering as inputs of layer M the outputs of layer M-1. Alternatively, we can maintain the transformation coefficients of layer M and act on one of the layers of the preceding levels. It is fundamental that the manager act on only a single gi, choosing the one which eliminates the error, no matter the layer it pertains to.

(b) *Composite, or multi-lever, strategies*. The manager can act on several levers at the same time, trying to simultaneously modify several transformation coefficients (action rates). A composite strategy can be identified fairly easily by operating on a chain or a path, that is, on an uninterrupted sequence of arrows which, starting from any input Xh, reaches Y. Setting the final output as $Y = Y^*$ and proceeding backwards, the manager tries to identify for modification one or more Yh from the preceding layer in order for Y to assume the value Y^*. Values for the "g_h" coefficients (if admissible) are sought which allow Y to be attained at the end of the causal chain.

(c) *Global strategies*, which seek to modify the values of all the transformation rates, g_i, for all the variables that precede Y. It is difficult for the manager to make perfect calculations outside of extremely simply situations.

A simple global strategy (usually used in neural networks without direct feedback) is to effect a random change of the g_i in the following manner:

1. Assign rectifiers, h_i, to all the g_i coefficients whose values are randomly (or intuitively) set by the control system manager.
2. Calculate the new value for Y1; if the error increases, assign new correctors, again randomly (or intuitively) chosen.
3. Repeat these operations until $E(Y)$ is eliminated or the error cannot be further reduced.

A general strategy which can always be applied, which is a simplification of the above strategy, is shown in the general model in Fig. 10.12. This model recalculates, for each layer, the coefficients needed to achieve the objective, $Y* = 150$.

Assigning to the Xi the input values shown in the cells in the first row ($X_1 = 20$, $X_2 = 10$, and $X_3 = 5$) we follow the arrows that indicate the causal links that produce the transformations. Each cell next to the arrows indicates the theoretical values of the "old g_i" coefficients, which allow the values of the "old Y" in the various layers to be calculated through the values of X. Using the "old g_i," the "old Y" value output of 270 is obtained. Since the objective has been set at $Y* = 150$, the error is equal to the "old $E(Y)$" or -120. Therefore, it is necessary to again determine the "old g" values for all the transformations in order to find a correct "new Y" value equal to 150, which produces a "new $E(Y)$" = 0. The *automatic controller* in the lower part of Fig. 10.12 enables the "new g_i" to be calculated (following the strategy which is apparent by examining the figure)—which produces the "new Y" in the various layers—and the "old $E(Y)$" to be eliminated, thus achieving the objective and maintaining stable the initial input values of X.

The model in Fig. 10.12 is flexible since it even allows the system manager to adapt various coefficients, thereby facilitating his or her task. In fact, when an "original" coefficient in the model is replaced by a "new" one, all the old "g_i" are automatically recalculated, so that the manager can then proceed to the recalculations based on any hypothesis of change whatsoever, choosing the alternative he or she holds to be most effective and to represent the preferred strategy. It is possible to modify a single transformation coefficient or several coefficients. For each choice, the model recalculates all the coefficients. Obviously the *automatic calculator* is programmed specifically for the transformation system in Fig. 10.12, but the principle can also be applied to more complex structures.

In order to choose the best strategy, a cost/benefit analysis is required, as described in Sect. 4.7. In fact, rectifying the "old g_i" coefficients likely will entail a cost for the *technical* intervention on some apparatus, thus necessitating rectifications which are more efficient in terms of both implementation time and cost.

In order to clarify the above concepts, let us assume that the model in Fig. 10.12 represents a hydraulic network which, receiving liquid at various amounts of pressure from three sources ($X_1 = 20$, $XX_2 = 10$, and $X_3 = 5$), varies, at successive stages, the overall pressure based on the transformation coefficients. The value Y indicates the pressure obtained, while $Y*$ indicates the desired pressure. If the
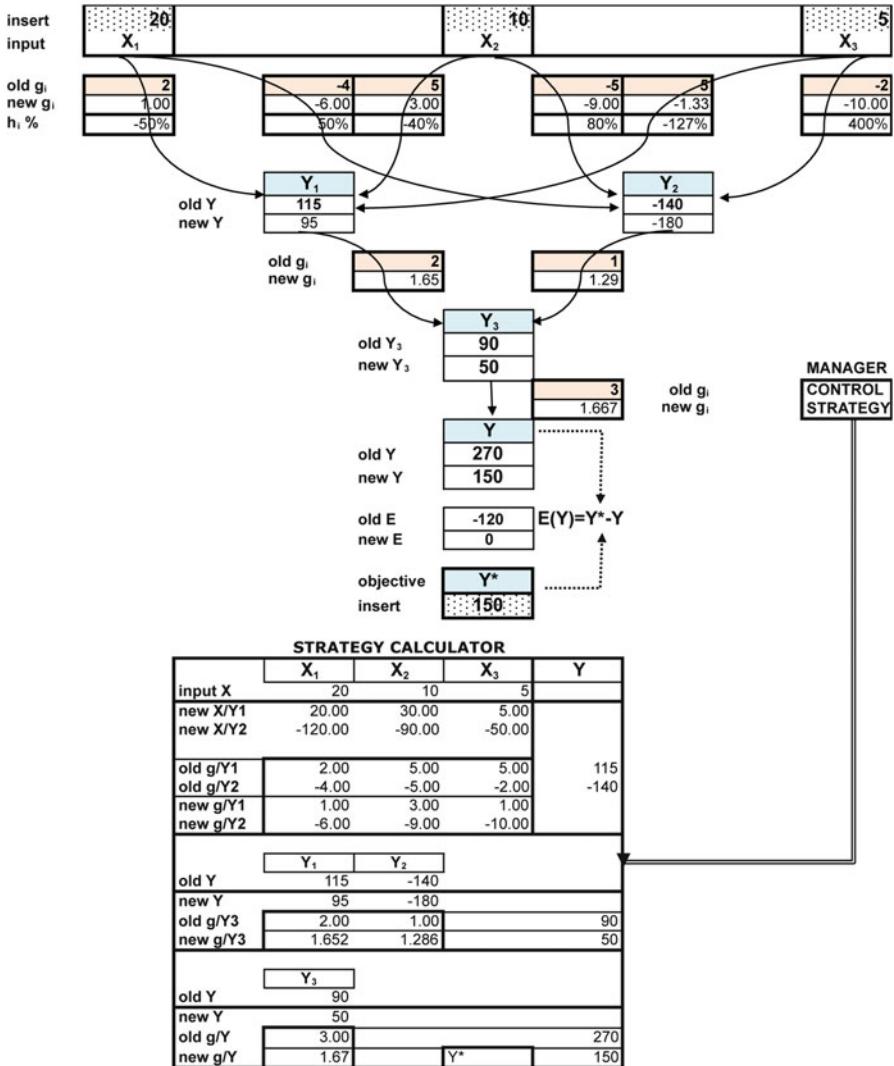
Fig. 10.12 Control of a multi-layer transformation system

incoming pressure is not varied, which is shown alongside the X_i, it becomes necessary to modify the outgoing pressure by means of appropriate pumps and valves, shown alongside each Y_i preceding Y (taking into account the various layers). We can intuitively understand from this example that if the manager were "near" the outflowing liquid, Y, then the control of the system, achieved by changing the first-level flow coefficients (from X to Y_1 and Y_2), would entail a high cost, since the control values are numerous and far from the operator. In this case it would be simpler to adjust only those valves that regulate the coefficients on

the curved line that connects Y_3 to Y. Nevertheless, if this control should require a non-admissible variation of the installed regulator and it were necessary to substitute this with a more powerful regulator, with high energy consumption, then the steep cost of this strategy would make it more efficient to regulate the mechanisms located "near" the inputs.

Thus, we can say in general that the most effective strategy is the one that minimizes the cost of the control in terms of both time and money, assuming equal advantages in reducing $E(Y)$.

10.5 Summary

In this final chapter I have presented some concluding considerations to stimulate further reflection:

1. In Sect. 10.1 I have outlined several fundamental general *hypotheses* in order to propose a control discipline.
2. In Sect. 10.2 I have discussed the *human aspects* of control.
3. In Sect. 10.3 I have used FAQs to investigate the content and limits of this book.
4. In Sect. 10.4 I have suggested that the Magic Rings also act in contexts we might never imagine:

    (a) In the control carried out by (in) religions and regarding life practices (Sect. 10.4.1).
    (b) In the control of the chaotic dynamics produced in our environment, as shown by the control of deterministic chaos (Sect. 10.4.2).
    (c) In the control of *variant structure* systems (Sect. 10.4.3), where the X variables cannot be modified and the control is achieved through appropriate values of the "g" and "h" rates. Three fundamental classes of variant action systems (which were mentioned only in passing) are (1) transformation systems, (2) production systems, and (3) neural networks.

I have only briefly discussed these types of control systems, which nevertheless deserve a broader analysis which goes beyond the scope of this book. Thus, I have only presented transformation systems and a general algorithm to illustrate the control strategies.

Glossary

Action rate "$g(Y/X)$" Indicates the variation of Y for each unit of variation of X.

Active variable, X_t In the general model of the control system, any variable X capable of modifying Y. See *Levers*.

Apparatuses See *Chain of control*.

Archetypes Models or patterns that occur continuously, also known as *generic* structures. "*One of the most important, and potentially most empowering, insights to come from the young field of systems thinking is that certain patterns of structure recur again and again. These "systems archetypes" or "generic structures" embody the key to learning to see structures in our personal and organizational lives*" (Senge 2006, p. 93).

Causal loop diagram (CLD) Model of a system built following systems thinking using arrows to indicate the direction of the causal relationships among the variables. A CLD consists of a system of loops in which all variables are linked by arrows, without there being an initial and a final variable. All the variables are connected. By connecting a number of variables and determining the direction of variation we can build models of every dynamic system, keeping in mind that we must zoom in order to analyze the processes in more detail, in order to identify and connect other important variables.

Chain of control The fundamental "machines" or "apparatuses" that represent *physical control systems* and whose functioning produces the dynamics in the active and passive variables and determines the variance.

- EFFECTOR: produces a variation in X_t into the corresponding variation in Y_t.
- DETECTOR (sensor or comparator): measures the value of Y, compares this with the objective Y^* (or the constraint Y°), and determines the deviation $E(Y)$.
- REGULATOR (or compensator): "activates" the lever X taking account of $E(Y)$.
- INFORMATION TRANSMISSION: represents the "real" *chain of control* that produces the "formal" control system.

Control apparatuses See *Chain of control*.

Control discipline This is based on the hypothesis that control systems occupy a preeminent place among all types of systems. Even though we are not

accustomed to "seeing them," they are all around us, and only their presence makes our world, life, society, and our very same existence possible, producing an ordered and livable world, erecting barriers against disorder, and guiding the dynamics of the system toward equilibrium states. The control discipline teaches us how to recognize control systems, build models, simulate their behavior, and use these to improve our knowledge and our behavior. I would propose naming the control discipline as the Sixth Discipline, or the *control discipline of the individual, the collectivity and the organizations in the ecosystem*; this is the discipline of the present and future of our world.

Control objective Each specific value, Y^*, or each trajectory of values, Y_t^*, however formed, which are set as values that the *objective variable*, Y_t, must achieve or maintain. The control objective can also mean the *constraint* which is not to be exceeded or the *limit* to be respected. If Y^* is a vector of values, $[Y^*]$, then the control system is multi-objective and the system manager must establish a *control policy*. The objectives can be both quantitative and qualitative.

Control process See *control system (concept)*.

Control system (concept) Any set of apparatuses, *logical* or *technical* (algorithm or machine, rule or structure, etc.), which, for a set of instants, perceives $E(Y)_t$, calculates and assigns the values X_t, and produces the appropriate Y_t to gradually eliminate, when possible, the error $E(Y)_t = Y^* - Y_{t^*}$ at instant t^*. The control system is repetitive and functions by means of action (X acts on Y) and reaction ($E(Y)$ acts on X through Y), activating a closed-loop or feedback control. With a certain number of iterations on the control lever, it tries to achieve the objective (goal-seeking systems) or to respect the constraints or limits (constraint-keeping systems) by gradually eliminating the deviation $E(Y)$. Control systems have a *logical structure* and a *technical structure*. The former consists of the logical relationships between variables, always arranged in one or more balancing loops in order to develop feedbacks between $[X]$ and $[Y]$; the latter is formed by the *chain of control*.
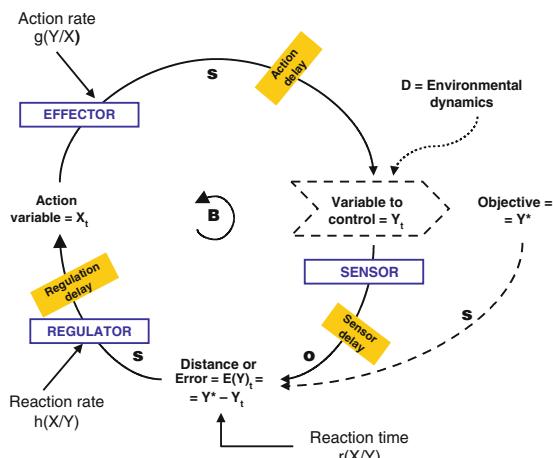
Control system (logical structure) The general model of a control system is a balancing loop, which consists of the elements shown in the model.

Control system (technical structure) The set of apparatuses that constitute the *physical control systems*. See *Chain of control*.

Control variable Synonym for *lever*.

Cybernetic control system An automatic control system guided toward a fixed (or even variable) objective determined by an outside governor, but guided by a manager within the chain of control.

Delay Abnormal or unexpected length of action of X on Y. There are three types of delay:

- *Action delay* (or response delay), which slows the response of Y to an impulse from X; this depends on the effector.
- *Detection delay* (or informational delay), which acts on our perception and on the measurement of the error.
- *Regulation* (decisional) *delay*, which occurs when the regulator does not respond promptly to the error.

Deviation See *Error, E*.

Direction of the causal link The arrow that indicates which variable is the *cause* and which the *effect* in a causal relation typical of systems thinking. The *cause* variable (X = input) is written at the tail of the arrow; the *effect* variable (Y = output) at the head (first arrow). If the causal link shows the opposite direction (second arrow), then X is the effect and Y the cause.

$$X \to Y$$
$$X \leftarrow Y$$

Direction of the variations This indicates the concordance in the signs of the variations between two variables, X and Y, which are linked in the same direction. Variations of X are linked to those of Y in the same direction (**s**) if Y increases/diminishes when X increases/diminishes (first arrow). Variations of X are linked to Y's variations in the opposite direction (**o**) if Y diminishes/increases when A increases/diminishes, thus presenting the opposite sign of variation (second arrow).

$$\underset{+/-}{X} \overset{\mathbf{s}}{\to} \underset{+/-}{Y} \qquad \underset{+/-}{X} \overset{\mathbf{o}}{\to} \underset{-/+}{Y}$$

Discipline Following Peter Senge: "A discipline is a developmental path for acquiring certain skills or competencies. [...] To practice a discipline is to be a lifelong learner. You "never arrive"; you spend your life mastering disciplines" (Senge 2006, p. 10).

Distance See *Error, E(Y)*.

Disturbances, D In the general model of the control system, each variable, D_t, external to the system, that alters the values of Y, regardless of the values of X. Disturbances can affect all the *apparatuses* that make up the *control chain*.

Error, $E(Y)$ (distance, deviation, gap, variance) In the general model of the control system, the variable "$E(Y)_t = \Delta(Y)_t = Y^* - Y_t$," which represents the *distance* or *deviation* between the values of the objective Y^* and those of "Y_t."

Feedback control systems Control system in which the control of Y_t is achieved through decisions to vary X_t over a succession of repetitions of the control cycle.

Feedforward control systems, also called *decision* or *one-shot control systems* Control system which try to reach the objective through a single cycle, and thus through a single *initial decision* made based on precise calculations; however, when an error is detected in the attempt to achieve Y^* the manager cannot correct this through other regulating decisions to eliminate the error. They are typically systems that "fire a single shot" to achieve the objective.

Forrester, Jay Founder of *system dynamics*, in his fundamental work *industrial dynamics* (1961).

Gap See *Error, E*.

Goals In the general model of the control system, any specific value, Y^*, or any trajectory of values, Y_t^*, however defined, that must be reached and possibly maintained by the control variables. If Y is a vector $[Y^*]$, then the control system is a multi-objective system and the manager must establish a *policy* of control. Goals can be quantitative or qualitative.

Governance of the control system The process through which a subject, the governor, determines the objective Y^* or the vector $[Y^*]$ that management must achieve.

Interfering control systems Two or more control systems interfere with each other if they are interrelated, so that the values of Y_A of a control system A determine and are determined by the values of Y_B of other interfering control systems. The interference can also occur in both directions.

iThink, Stella See *simulations softwares*.

I/O systems See *type of control systems*.

Levers In a control system, each control variable, X, which—through some process—acts on the *variable to control*, Y, in order to eliminate $E(Y)$. If X is a vector of variables $[X]$, then the system is multi-lever and the manager of the system should establish a *strategy* for the control.

Loop A closed causal chain formed by a circular link between two variables, X and Y, which can be linked by two opposite directions with respect to the causal link. Loops can be basic, when there are only two variables, or compound, when more than two variables are joined in a circular link. There are only two basic types of loop: *reinforcing loops* [**R**], which produce a reciprocal increase or reduction—in successive repetitions of the system's cycle—in the values of the two variables having reciprocal causal links; and *balancing loops* [**B**], which maintain relatively stable the values of the connected variables.
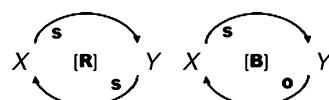
Manager of the control system The subject that maneuvers the control system in order to achieve the goals established by the governance, by adjusting the control levers and deciding on the *strategy* and *policy* to be adopted. The subject can be external to the control system or internal to the system itself, replacing some apparatus.

Multi-lever control systems See *type of control systems*.

Multi-objective control systems See *type of control systems*.

Objective variable, *Y* Synonym for *variable to control*.

ON/OFF systems See *type of control systems*.

Open causal chain Connection among a succession of several variables entailing some causal link.

$$X \xrightarrow{\text{s/o}} Y \xrightarrow{\text{s/o}} Z$$

Passive variable, Y_t Synonym for *variable to control*.

Period of control The number of intervals, over the typical time frame of the system, which are necessary to stably achieve the objective in the absence of outside *disturbances*.

Pluri-lever Synonym for *multi-lever*.

Pluri-objective Synonym for *multi-objective*.

Policy (in control) In multi-objective systems, policy represents the order of priority in achieving different objectives.

Powersim See *simulation softwares*.

Problem solving (PS) The process through which we seek the solution of any problem. The PS can be interpreted as a control system, because a problem can be interpreted as an error or deviation, $E(Y)$, between two states—current (Y) and desired (Y^*)—of a variable, Y (which we assess as useless or harmful) and that we want to eliminate or reduce as far as possible by acting on the variables X that can change Y. In PS, solving the problem means designing the control system in order to identify the lever X_t, which, by producing the values Y_t—taking into account the states of nature, D_t—enables you to eliminate $E(Y)$. Notice that the deviation is not the problem but the symptom that is revealed. The problem lies in the malfunctioning of the system that generates the symptom.

Reaction rate "$h(X/Y)$" Indicates the variation in X for each unit of variation in Y; usually "$h = 1/g$" if the system is perfectly symmetrical to the control.

Reaction time "$r(X/Y)$" (time to eliminate the error) Indicates the speed with which the control system moves toward the objective; a reaction time of $r = 1$ indicates an immediate but sudden control; a reaction time $r > 1$ makes the control slower but smoother.

Regulation In a broad sense, this is synonymous with control. In a strict sense, it is a process that assigns values to the *control levers*, based on the *deviation* and the *reaction rate*.

Senge, Peter He is considered to be the popularizer of systems thinking, with his fundamental text: *Fifth Discipline: The Art and Practice of the Learning Organization* (Senge, 1990).

Simulation softwares Software programs (simulation tools) created to simulate the behavior of systems with a high number of variables. I would mention:

- Powersim (www.powersim.com).
- MyStrategy and SYSDEA (www.strategydynamics.com/mystrategy/).
- ithink e stella (www.iseesystems.com/index.aspx).
- Vensim (www.vensim.com).
- Excel Software (www.excelsoftware.com/).

Sixth discipline See *control discipline*.

Stock and flow diagram (SFD) or level and flow structure (LFS) System dynamics theory assumes that a dynamic system can normally be viewed as composed of *stock* or *level* variables and of *flow* or *rate* variables that change the amount of stock. The flow variables that increase the stock (in various ways to be defined) can be considered as INPUTS; those that decrease it, as OUTPUTS. Several *simulation softwares* are based on this logic.

Strategy (in control) Control with two independent control variables *always* implies a *strategy* that defines an *order of priorities* regarding action on the control "levers."

Symptom In *problem solving*, understood as a process of control, this indicates the meaning that the manager gives to the deviation $E(Y)$.

System (from a systems thinking perspective) Systems thinking defines a "system" as a unitary set of interconnected variables possessing its own autonomy, which is capable of producing emerging macro-dynamics that do not coincide with any of the micro-dynamics of the individual variables or their partial subsystems.

System dynamics The technique for translating the qualitative models (causal loop diagrams) into quantitative (stock and flow diagrams) ones that, by quantifying the initial values of the temporal variables and specifying the variation parameters, are able to generate the dynamics of those variables. This technique was developed as industrial dynamics by *Jay Forrester* in the 1960s.

Systems archetypes See *Archetypes*.

Systems thinking Discipline that recommends that we observe the world as a system of dynamic systems, recursive and repetitive—often with memory— each composed of interacting and interconnected variables. Systems thinking was presented by Peter Senge in The Fifth Discipline: The Art and Practice of the Learning Organization (Senge, 1990). Systems thinking does not only represent a specific technique for constructing models but also a mental attitude, an approach, a logic, and a language. The systems are represented through qualitative models called *causal loop diagrams*.

Systems thinking laws There are two fundamental laws that derive from systems theory:

1. *The Law of Structure and Component Interaction*: in order to understand and control the dynamics in the world it is necessary to identify the systemic structures that make up this world. Corollary: in observing a dynamic world, the "*ceteris paribus*" assumption is never valid.

2. *Law of Dynamic Instability*: expansion and equilibrium are processes that do not last forever; they are not propagated ad infinitum. Sooner or later stability is disturbed. Sooner or later the dynamics are stabilized.

Type of control systems (CSs) Minimal typology:

- *Artificial* and *natural* CSs; Control systems designed and constructed by man are artificial or manmade control systems.
- *Manually controlled* systems, *automatic* CSs, or *cybernetic* systems. Systems which involve continuous manual control by a human operator are manually controlled control systems; control systems which incorporate the manager in their chain of control can be defined as automatic control systems.
- *Quantitative* CSs; the control is *quantitative* if the variable Y_t represents a *quantitative measure* and the objective is a value that Y_t must reach (*set reference*) or a set of values to be achieved (*track reference*).
- *Qualitative* CSs; systems in which both the variable Y_t and the objective Y^* are qualitative in the broadest sense of the term (objects, colors, flavors, forms, etc.).
- *Attainment* CSs; I shall define *attainment control systems* as those control systems that act to "attain" a *quantitative* objective, independently of the control context.
- *Recognition* and *identification* CSs; a general class of qualitative control system where the qualitative variable to control, Y_n, represents an "object" that the system *scans* in order to recognize or identify a "model-objective," Y^*.
- *Steering* CSs (also called "two-way" systems); systems acting to achieve the objective through positive or negative adjustments of Y_t that converge toward Y^*, independently of the initial value of Y.
- *Halt* CSs (also called "one-way" systems); systems in which Y^* is by nature a non-exceedable limit or constraint.
- *Collision*, *anti-collision*, and *alignment* CSs. Two interfering systems whose objectives are dynamic and tend to coincide are defined as "collision systems." If the objectives tend *not* to coincide they are defined as "anti-collision systems." If the objectives maintain a constant distance they are defined as "alignment systems."
- *Goal-seeking* CSs. Those in which Y^* represents a quantitative objective: to achieve a result, or to attain or maintain operating standards.
- *Constraint-keeping* CSs, in which Y^* represents a constraint to respect or a limit not to be exceeded; these are usually halt systems.
- *Regulator* CSs, where Y^* (objective or constraint) is a specific value (goal, constraint, or limit) that Y_t must reach and maintain over time.
- *Tracking* CSs, or *path systems*, where Y^* (goal, constraint, or limit) is a *trajectory*, that is, a sequence of values, Y^*_t, however formed, which Y_t must follow.
- *Multi-lever* or pluri-lever CSs, if the control is via a vector of levers $[X]$; if $[X] = X$, the system is single lever.
- *Independent levers* CSs; multi-lever systems whose levers can be adjusted independently; dependent lever systems, if the levers allow only variations in the opposite direction.
- *Multi-objective*, or pluri-objective CSs; if the control is a vector of objectives $[Y^*]$; if $[Y]^* = Y^*$, the system is single objective, in which case $[Y^*]$ must be equivalent to $[Y]$.
- *Independent objectives* CSs; multi-objective systems whose objectives can be achieved independently; bound objective systems if a goal impedes the achievement of other goals.

- CSs with or without *delays*.
- CSs of *direct* or *indirect control*; a control system is direct if X and Y vary in the same direction "**s**"; otherwise, it is indirect.
- *Autonomous*, or *interfering* CSs; interfering systems are those whose values of Y are influencing each other.
- CSs with *fixed* or *variable objectives* (or systems of pursuit); in the former, the variable must achieve a passive constant, Y^*; in the latter, the target Y_t^* represents a variable that depends on "t," Y and/or on X;
- *On/off* CSs. Halt systems that achieve Y^* by activating [on] the X_t lever for a set time, T^*, until the system stops [off], only to start up again when the disturbances, D, once again produce an error of a set amount, ΔE^*.
- *I/O* CSs. Particular *steering* systems that try to achieve the objective Y^* by "turning on" [I] the X lever once or several times and for a given length of time, which is decided on each occasion by the manager, in order to obtain a fixed value ΔY, after which the lever is "turned off" [O], thereby eliminating $E(Y)$. The lever is once again turned on when D produces another error.
- *Tendential* CSs. In these systems Y_t represents a time series (for example, the scores over time in an archery competition by participants trying to improve their performance) which, through some statistical procedure (trend, moving average, etc.), aims to ensure that Y_t tends toward a value-objective, Y^*, through a statistical measure.
- *Combinatory* CSs. Systems in which each Y_t represents a synthetic value, $[Y_{tN}]$, derived from the combination (to be specified in some manner) of the values of a vector of N variables, each of which is controlled by its own control variable.

Variable to control Any variable Y controlled by X in order to achieve a value Y^* set as an objective.

Symbols (minimal)

X_t Values of the control lever
Y_t Values of the variable to be controlled
Y^* Fixed objective
Y_t^* Variable objective
D_t External disturbance values
$\Delta(Y)_t = Y^* - Y_t$ (or $E(Y)_t$) Deviation, gap, error for systems [$\mathbf{s} \rightarrow \mathbf{o} \rightarrow \mathbf{s}$]
$\Delta(Y^*)_t = Y_t - Y^*$ (or $E(Y^*)_t$) Deviation, gap, error for systems [$\mathbf{s} \rightarrow \mathbf{s} \rightarrow \mathbf{o}$]
$g(Y/X)$ Action rate
$h(X/Y)$ Reaction rate
$r(X,Y)$ Reaction time

References[1]

Ackoff, R. L., & Sasieni, M. V. (1968). *Fundamentals of operations research*. London, UK: Wiley.

Adams, W. M. (2006). *The future of sustainability: Re-thinking environment and development in the twenty-first century*. http://iucn.org

Ahmed, E., Elgazzar, A. S., & Hegazi, A. S. (2005). *An overview of complex adaptive systems*. arXiv:nlin/0506059 [nlin.AO]. http://arxiv.org/abs/nlin/0506059v1

Allen, P. M. (1997). *Cities & regions as self-organizing systems: Model of complexity*. Amsterdam, Netherlands: Gordon and Breach.

Alleyne, R. (2010, September 21). Titanic sunk by steering blunder, new book claims. *The Telegraph*. www.telegraph.co.uk/culture/books/booknews/8016752/Titanic-sunk-by-steering-blunder-new-book-claims.html

Anderson, V., & Johnson, L. (1997). *Systems thinking basics. From concepts to causal loops*. Waltham, MA: Pegasus Communications.

Ankersmit, F. R., & Mooij, J. J. A. (1993). *Knowledge and language* (Metaphor and knowledge, Vol. III). Dordrecht, The Netherlands: Kluwer.

Anthony, R. N. (1965). *Planning and control systems: Framework for analysis*. Boston, MA: Harvard University.

Anthony, R. N. (1988). *The management control functions*. Boston, MA: HBS Press.

Arbib, M. A. (1987). *Feedback and feedforward*. http://www.answers.com/topic/feedback-and-feedforward

Argyris, C. (1993). *Knowledge for Action. A guide to overcoming barriers to organizational change*. San Francisco, CA: Jossey Bass.

Argyris, C., & Schön, D. (1974). *Theory in practice. Increasing professional effectiveness*. San Francisco, CA: Jossey Bass.

Arnold, G., & Davies, M. (2000). *Value-based management: Context and application*. London, UK: Wiley.

Arsham, H. (2006). Economic order quantity and economic production quantity. *Models for Inventory Management*. http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/Inventory.htm

---

[1] We know nothing at all. All our knowledge is but the knowledge of schoolchildren. The real nature of things we shall never know (Albert Einstein).

As this book deals with Control Systems considered from various points of view, a complete bibliography may appear to be boundless. Here I have only listed the works cited in the book. All the sites mentioned have been visited in April 2014. In order to make it easier for the reader to directly examine the sources, I have tried, whenever possible, to indicate those sources available online.

Arthur, W. B. (1973). *Optimal control theory with time delay*. Ph.D. Thesis, University of California, Berkeley.

Arthur, W. B., Durlauf, S. N., & Lane, D. A. (1997). *The economy as an evolving complex system II*. Reading, MA: Addison-Wesley.

Ashby, R. W. (1957). *An introduction to cybernetics* (2nd ed.). London, UK: Chapman & Hall.

Austin, J. L. (1962). *How to do things with words*. Cambridge, MA: Harvard University Press.

Bakshi, U. A., & Bakshi, V. U. (2009). *Automatic control systems*. Pune, India: Technical Publications.

Banathy, B. (2000). *Guided evolution of society. A systems view*. New York, NY: Springer.

Band, W. (1991). *Creating value for customers: Designing and implementing a total corporate strategy*. London, UK: Wiley.

Bampi, F. Retrieved from http://www.francobampi.it/franco/ditutto/cattolica/formule.htm

Barabási, A. (2002). *Linked: The new science of networks*. Cambridge, MA: Perseus.

Barrows, E. M. (2001). *Animal behavior desk reference: A dictionary of animal behavior, ecology and evolution*. Boca Raton, FL: CRC Press.

Bateson, G. (1972). *Steps to an ecology of mind* (1st ed.). Chicago, IL: University of Chicago Press. Excerpts, Google Books.

Bateson, G. (1979). *Mind and nature: A necessary unity* (1st ed.). (Advances in systems theory, complexity, and the human sciences). Cresskill, NJ: Hampton Press. (Originally published by Bantam, 1979). Excerpts, http://www.oikos.org/mind&nature.htm

Bateson, G. (2000). *Steps to an ecology of mind*. Chicago, IL: University of Chicago Press. Excerpts, Google Books.

Bateson, G. (2002). *Mind and nature: A necessary unity*. (Advances in systems theory, complexity, and the human sciences). Cresskill, NJ: Hampton Press. (Originally published by Bantam, 1979). Excerpts, http://www.oikos.org/mind&nature.htm

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science, 16*(5), 351–355.

BCG, The Boston Consulting Group. (1970). Corporate portfolio management: Theory and practice. *Applied Corporate Finance, 23*(1), 63–76.

Bednarz, J. (1988). Autopoiesis: The organizational closure of social systems. *Systems Research, 5*(1), 57–64.

Beer, S. (1966). *Decision and control*. London, UK: Wiley.

Beer, S. (1973). Fanfare for effective freedom. In J. Glassman & A. Chandra (Eds.) (2006) (pp. 52–69).

Beer, S. (1979). *The heart of enterprise*. London, UK: Wiley.

Beer, S. (1981). *Brain of the firm* (2nd ed.). London, UK: Wiley.

Beer, S. (1992). Preface. In H. R. Maturana & F. J. Varela (1992).

Beer, S. (1995). *Diagnosing the system for organizations*. London, UK: Wiley.

Berkeley University of California. (2010). *Guide to managing human resources*. http://hrweb.berkeley.edu/guide/performance.htm#top

Bhatawdekar, S. (webpage). http://quality-circles.blogspot.it

Blackwell, D., & Girshich, M. A. (1954). *Theory of games and statistical decisions*. London, UK: Wiley.

Bodhi, B. (1999). The Noble Eightfold Path. *The way to the end of suffering*. Retrieved from www.accesstoinsight.org/lib/authors/bodhi/waytoend.html

Bostic-St. Clair, C., & Grinder, J. (2001). *Whispering in the wind*. Scotts Valley, CA: J & C Enterprises.

Boulding, K. E. (1956). General system theory, management science. The skeleton of science. *Management Science, 2*(3), 197–208.

Bouyssou, D., Marchant, T., Pirlot, M., Tsoukias, A., & Vincke, P. (2006). *Evaluation and decision models with multiple criteria*. New York, NY: Springer Sciences + Business Media.

Braithwaite, R. B. (1953). Scientific explanation. A study of the function of theory, probability and law in science. *Tarner Lectures*. CUP Archive, Cambridge University Press.

Breventano, S. (1570). A history of antiquity, nobility, and of notable things of the city of Pavia (original: Istoria della antichità, nobiltà, et delle cose notabili della città di Pavia). Pavia: Hieronimo Bartholi, MDLXX. Retrieved from, http://books.google.it/books?id=_CQ8AAAAcAAJ&printsec=frontcov-er&hl=it&source=gbs_ge_summary_r&cad=0#v=onepage&q=torri&f=false

Britt, R. R. (2004). *Asteroid with chance of hitting earth in 2029 now being watched 'very carefully'*. www.space.com/scienceastronomy/asteroid_risk_041224.html

Brun, H. (1970). *Technology and the composer*. http://www.herbertbrun.org/techcomp.html

Brundtland Commission. (1987). *Our common future*. United Nations General Assembly. World Commission on Environment and Development (WCED)—The "Brundtland Commission".

Buchanan, K. L. (2000). Stress and the evolution of condition-dependent signals. *Trends in Ecology & Evolution, 15*(4), 156–160.

Buswell, R. E., & Lopez, D. S. (2003). *The Princeton dictionary of Buddhism*. Princeton, NJ: Princeton University Press.

Capra, F. (1996a). The rise of systems thinking. In F. Capra (Ed.), *The Web of life*. http://www.maaber.org/sixth_issue/deep_ecology_1e.htm#_ftnref35

Capra, F. (1996b). *The Web of life: A new scientific understanding of living systems*. New York, NY: Anchor Books.

Caprara, G. (2010, July 1). Nuvola bianca, porta pioggia [White cloud brings rain]. *Corriere della Sera*. http://www.corriere.it/scienze/10_luglio_01/nuvole-bianche_93fb9560-8537-11df-a3c2-00144f02aabe.shtml

Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York, NY: Springer.

Casey, W. (2000). *Project tools for leading organizational change*. Wheat Ridge, CO: Executive Leadership Group.

CASG, Complex Adaptive Systems Group. (2013). http://www.cas-group.net

Castelfranchi, C. (1998). Through the minds of the agents. *Journal of Artificial Societies and Social Simulation*, *1*(1). http://jasss.soc.surrey.ac.uk/1/1/5.html

Casti, J. L. (1985). *Nonlinear system theory* (Mathematics in science and engineering). Orlando, FL: Academic Press.

Ceccato, S. (1969). *Course in operational linguistic [Italian: Corso di linguistica operativa]*. Milano, Italy: Longanesi.

Champernowne, D. G. (1969). *Uncertainty and estimations in economics*. Edinburg, TX: Oliver e Boyd.

Chandler, A. D., Jr. (1962). *Strategy and structure: Chapters in the history of the American industrial enterprise*. Cambridge, MA: MIT Press.

Checkland, P. (1999). *Systems thinking, systems practice*. London, UK: Wiley.

Christen, M., & Schmidt, S. (2012). A formal framework for conceptions of sustainability—A theoretical contribution to the discourse in sustainable development. *Sustainable Development, 20*(6), 400–410.

Churchman, C. W., & Ackoff, R. L. (1954). An approximate measure of value. *Operations Research, 2*, 172–187.

Churchman, C. W., Ackoff, R. L., & Arnoff, E. L. (1957). *Introduction to operations research*. London, UK: Wiley.

Clark, A. (1989). *Microcognition: Philosophy, cognitive science, and parallel distributed processing*. Cambridge, MA: MIT Press. http://www.philosophy.ed.ac.uk/people/clark/pubs/microfx.pdf

Conte, R., & Paolucci, M. (2001). Intelligent social learning. *JASSS Journal of Artificial Societies and Social Simulation*. http://jasss.soc.surrey.ac.uk/4/1/3.html

Cooper, R. (1989). The rise of activity based costing—Part three: How many cost drivers do you need and how do you select them? *Journal of Cost Management, 2*(4), 34–46.

Cooper, R., & Kaplan, S. (1991, May–June). Profit priorities from activity-based costing. *Harvard Business Review*, 130–135.

Copeland, T., Koller, T., & Murrin, J. (2000). *Valuation: Measuring and managing the value of companies* (3rd ed.). London, UK: Wiley (2nd ed., 1996).

Copi, I. M., & Cohen, C. (2011). *Introduction to logic* (14th ed.). Edinburgh, TX: Pearson.

Costanza, R., & Patten, B. C. (1995). Defining and predicting sustainability. *Ecological Economics, 15*(3), 193–196.

Coyle, R. G. (1977). *Management system dynamics*. London, UK: Wiley.

Crosby, P. (1979). *Quality is free*. Boston, MA: McGraw-Hill.

Crow, K. (2002). *Value analysis and function analysis system technique*. http://www.npd-solutions.com/va.html

CSS, Complex Systems Society. (2013). *About complex systems*. http://www.complexssociety.eu/aboutComplexSystems.html

Cupples, B. (1977). Three types of explanation. *Philosophy of Science, 44*(3), 387–408.

Cyert, R. M., & March, J. G. (1963). *A behavioral theory of the firm* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Daniels, A. C., & Daniels, J. E. (2004). *Performance management: Changing behavior that drives organizational effectiveness* (4th ed.). Atlanta, GA: Performance Management Publications (1st ed., 1992).

Darwin, C. (1859). *The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. London, UK: Murray.

Davenport, T. H., & Stoddard, D. B. (1994). Reengineering: Business change of mythic proportions? *MIS Quarterly, 18*(2), 121–127.

Davis, K. (1949). *Human society*. New York, NY: Macmillan.

Dawkins, R. (1989). *The selfish gene*. Oxford, UK: Oxford University Press (1st ed., 1976).

Dawkins, R. (2004). *The ancestor's tale: A pilgrimage to the dawn of evolution*. Boston, MA: Houghton Mifflin.

Dawkins, R. (2006). *The God delusion*. Boston, MA: Houghton Mifflin Co.

de Carvalho, J. L. (1993). *Dynamical systems and automatic control*. Englewood Cliffs, NJ: Prentice Hall.

de Geus, A. (1988). Planning as learning. *Harvard Business Review, 66*(2), 70–74.

de Geus, A. (1992). Modelling to predict or to learn? *European Journal of Operational Research, 59*(1), 1–4.

de Geus, A. (1997). *The living company. Growth, learning and longevity in business*. London, UK: Nicholas Brealey.

de Geus, A. (2002). *The living company: Habits for survival in a turbulent business*. Boston, MA: Harvard Business Review Press.

de Saussure, F. (1916). *Course in general linguistics* (Original: Cours de linguistique générale). Charles Bally and Albert Sechehaye. New York: Philosophical Library, 1959. Excerpt, http://faculty.smu.edu/dfoster/cf3324/saussure.htm

Demartini, C. (2013). *Performance management systems. Design, diagnosis and use*. New York, NY: Springer.

Deming, W. E. (1982). *Out of the crisis*. Cambridge, MA: MIT.

Dennett, D. C. (1995). *Darwin's dangerous idea. Evolution and the meanings of life*. New York, NY: Simon & Schuster.

Dermer, J. D., & Lucas, R. G. (1986). The illusion of managerial control. *Accounting, Organizations and Society, 11*(6), 471–482.

Dewan, E. M. (1969). *Cybernetics and the management of large systems*. New York, NY: Spartan Books.

Drucker, P. F. (1986). *Management. Tasks, responsibilities, practices*. New York, NY: Truman Talley Books.

Dunlap, J. C., Loros, J., & De Coursey, P. J. (2003). *Chronobiology: Biological timekeeping*. Sunderland, MA: Sinauer.

Dyson, F. (1988). *Infinite in all directions*. New York, NY: Harper and Row.

Ebeling, C. E. (1997). *An introduction to reliability and maintainability engineering*. Boston, MA: McGraw-Hill.

Eccles, J. C., & Robinson, D. N. (1984). *The wonder of being human. Our brain & our mind*. New York, NY: Free Press.

Eco, U. (1976). *A theory of semiotics*. Bloomington, IN: Indiana University Press.

Edelman, G. (1989). *The remembered present by a biological theory of consciousness (review by Bobby Matherne 1997)*. New York, NY: Basic Books.

Edelman, G. (1992). *Bright air, brilliant fire: on the matter of the mind*. New York, NY: Basic Books–Harper Collins.

Edwards, M. (2003). *A brief history of Holons*. http://www.integralworld.net/edwards13.html

Ellison, B., Fisher, D. A., Linger, R. C., Lipson, H. F., Longstaff, T., & Mead, N. R. (1997). *Survivable network systems: An emerging discipline*. Survivable Network Technology Team CERT® (revised May 1999).

Environment. (2013). http://www.environment.gen.tr/gaia/70-gaia-hypothesis.html

Ethnologue, Languages of the World. (2013). http://www.ethnologue.com/ethno_docs/distribution.asp?by = area

Ezechiel, M. (1938). The cobweb theorem. *Quarterly Journal of Economics, 52*, 255–280.

Fayol, H. (1949). *Administration Industrielle et Générale [General and industrial management]*. New York, NY: Pitman Publishing Corporation (Original work published in 1916).

Ferreira, A., & Otley, D. (2009). The design and use of performance management systems: An extended framework for analysis. *Management Accounting Research, 20*(4), 263–282.

Fishburn, P. C. (1964). *Decision and value theory*. London, UK: Wiley.

Fishburn, P. C. (1967). Methods of estimating additive utilities. *Management Science, 13*, 435–453.

Fisher, D. A., & Lipson, H. F. (1999). Emergent algorithms—A new method for enhancing survivability in unbounded systems. *The 32nd Hawaii International Conference on System Sciences*. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.8956&rep=rep1&type=pdf

Flake, G. W. (2001). *The computational beauty of nature*. Cambridge, MA: The MIT Press.

Flamholtz, E. G. (1996). *Effective management control: theory and practice*. New York, NY: Springer.

Fleming, J. R. (2007). The climate engineers. *Wilson Quarterly*. http://www.colby.edu/sts/climateengineers.pdf

Flood, R. L. (1999). *Rethinking the fifth discipline*. New York, NY: Routledge.

Forrester, J. W. (1961). *Industrial dynamics*. Waltham, MA: Pegasus Communications.

Forrester, J. W. (1971). *Counterintuitive behavior of social systems* (Collected papers of J. W. Forrester, pp. 211–244). Cambridge, MA: Wright-Allen Press.

Forrester, J. W. (1983). *Foreward*. In N. Roberts, D. Andersen, M. S. Garet, W. A. Shaffer, R. M. Deal

Forrester, J. W. (1991). *System dynamics and the lessons of 35 years*. Sloan School of Management, M.I.T. ftp://nyesgreenvalleyfarm.com/documents/sdintro/D-4224-4.pdf

Forrester, J. W. (1999). *System dynamics: The foundation under systems thinking*. Cambridge, MA: Sloan School of Management, MIT 02139. http://clexchange.org/ftp/documents/system-dynamics/SD2011-01SDFoundationunderST.pdf

Gell-Mann, M. (1992). Complexity and complex adaptive systems. In Hawkins & Gell-Mann (Eds.) (1992). http://tuvalu.santafe.edu/~mgm/Site/Publications_files/MGM%20107.pdf

Gell-Mann, M. (1994). *The quark and the jaguar* (p. 1994). New York, NY: W. H. Freeman.

Gell-Mann, M. (1995). What is complexity? *Complexity, 1*(1), 16–19.

Gephart, M. A., Marsick, V. J., Van Buren, M. E., & Spiro, M. S. (1996). Learning organizations come alive. *Training & Development, 50*(12), 35–45.

Gibbon, D. (1998). Procedural explanations. Retrieved from http://coral.lili.uni-bielefeld.de/Classes/Summer98/PragEngDialogue/pragengdialogue/node18.html

Gilbert, N. (1995). Simulation: an emergent perspective. *New Technologies in the Social Sciences*. http://cress.soc.surrey.ac.uk/resources/emergent.html

Gilbert, N., & Troitzsch, K. G. (1999). *Simulation for the social scientist*. Buckingham, UK: Open University Press.

Gleick, J. (2008). *Chaos: Making a new science (Revised edition)*. London, UK: Penguin.

Glassman, J., & Chandra, A. (2006). Inventing systems. *Spring Quarter 2006*. http://ada.evergreen.edu/~arunc/texts/InventingSystems/readings.pdf

Goodwin, G. C., Graebe, S. F., & Salgado, M. E. (2001). *Control system design*. Englewood Cliffs, NJ: Prentice Hall.

Google, Driverless car project. (2012). http://en.wikipedia.org/wiki/Google_driverless_car

Google, Prius. (2010). http://www.autoblog.com/2010/10/14/video-abc-news-gets-taken-for-a-spin-in-googles-self-driving-t/

Gopal, M. (2002). *Control systems: Principles and design*. New Delhi, India: Tata McGraw-Hill Education.

Green, S. G., & Welsh, M. A. (1988). Cybernetics and dependence: reframing the control concept. *Academy of Management Review, 13*, 287–301.

Gupta, R. C. (2001). *Statistical quality control* (6th ed.). Delhi, India: Romesh Chader Khanna.

Habib, C. (2011). *As above, so below: The worldview of Lynn Margulis*. http://realitysandwich.com/128651/above_so_below_worldview_lynn_margulis/

Haines, S. (2005). *Enterprise-wide change: Superior results through systems thinking*. San Francisco, CA: Pfeiffer-John Wiley.

Håkansson, H., & Snehota, I. (1995). *Developing relationships in business networks*. London: Routledge.

Hammer, M., & Champy, J. (1993). *Reengineering the corporation*. New York, NY: Harper Business.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243–1248. http://dieoff.org/page95.htm

Hardin, G. (1985). *An ecolate view of the human predicament*. www.garretthardinsociety.org/articles/art_ecolate_view_human_predicament.html

Harris, R. A. (1998). *Introduction to decision making*. www.virtualsalt.com/crebook5.htm

Harris, R. A. (2002). *Creative problem solving: a step-by-step approach*. Los Angeles, CA: Pyrczak Publishing.

Hawking, S. W., & Mlodinow, L. (2010). *The grand design*. Bantam. Excerpt, http://bookre.org/reader?file=65691

Hawkins, J. A., & Gell-Mann, M. (Eds.). (1992). *The evolution of human languages* (Santa Fe Institute Series). Reading, MA: Addison-Wesley.

Hayes, R. B., & Jaikumar, R. (1988). Manufacturing's crisis: New technologies, obsolete organizations. *Harvard Business Review, 66*(5), 77–85.

Hecht, J. (2007). NASA analysis of asteroid risk deeply flawed, critics say, NewScientist.com news service. http://space.newscientist.com/article/dn11901-nasa-analysis-of-asteroid-risk-deeply-flawed-critics-say.html

Hedberg, B. (1981). How organizations learn and unlearn. In P. Nystrom & W. Starbuck (Eds.), *Handbook of organizational design* (pp. 3–28). New York, NY: Oxford University Press.

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays*. New York, NY: The Free Press.

Hempel, C. G. & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15(2)*, 135–175. Retrieved from http://www.jstor.org/discover/10.2307/185169?uid=3738296&uid=2&uid=4&sid=21103619160337

Heylighen, F. (2013). *The global superorganism: An evolutionary-cybernetic model of the emerging network society*. www.txtshr.com/computers/11212/index.html

Heylighen, F., & Joslyn, C. (2001). The law of requisite variety. *Principia Cybernetica*. http://pespmc1.vub.ac.be/REQVAR.html

Hiatt, J. M. (2006). *ADKAR: A model for change in business, government and our community*. Loveland, CO: Prosci Research.

Hofstede, G. (1978). The poverty of management control philosophy. *Academy of Management Review, 3*, 450–461.

Holland, J. H. (1995). *Hidden order*. Cambridge, MA: Perseus Books.

Hölldobler, B., & Wilson, E. O. (1990). *The ants*. New York, NY: Springer.

Houghton, J. T. (1994). *Global warming: the complete briefing* (1st ed.). Cambridge, UK: Cambridge University Press.

Hubbell, W. W., Jr. (1996). Combining economic value added and activity-based management. *Journal of Cost Management, 10*(1), 18–29.

Huxley, T. H. (1874). On the hypothesis that animals are automata, and its history. *Fortnightly Review, 16*, 555–580.

IEP, Internet Encyclopedia of Philosophy (2014). Theories of explanation. Retrieved from http://www.iep.utm.edu/explanat/

Imboden, D. M., & Pfenninger, S. (2013). *Introduction to systems analysis*. New York, NY: Springer.

Ishii, D. K., & Kline, S. L. (2008). Procedural explanations in mathematics writing. A framework for understanding college students' effective communication practices. *Written Communication, 25*(4), 441–461.

Ishikawa, K. (1980). *QC circle koryo: general principles of the QC circle*. Tokyo, Japan: Union of Japanese Scientists and Engineers (1st ed., 1970).

Jakobson, R. (1970). *Essais de linguistique générale* [*Course in general linguistics*] (1st ed. 1963). Les Editions de Minuit. http://nedelcu.ca/documents/Jakobson-Theo-de-la-communication.pdf

Johnson, R. A., Kast, F. E., & Rosenweig, J. E. (1963). *The theory and management of systems*. New York, NY: McGraw-Hill.

Johnson, G., & Scholes, K. (2002). *Exploring corporate strategy* (6th ed.). London, UK: Prentice Hall.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N., & Khemlani, S. (2013). Toward a unified theory of reasoning. In B. Ross (Ed.). *The psychology of learning and motivation* (pp. 1–42). Elsevier. http://mentalmodels.princeton.edu/papers/2013unified-theory.pdf

Kahn, A. E. (1988). *The economics of regulation: Principles and institutions*. Cambridge, MA: The MIT Press.

Kaplan, R. S., & Norton, D. P. (1992, February). The balanced scorecard. Measures that drive performance. *Harvard Business Review*, 71–79.

Kaplan, R. S., & Norton, D. P. (1996). *The balanced scorecard. Translating strategy into action*. Boston, MA: Harvard Business School Press.

Kaplan, R. S., & Norton, D. P. (2001). Transforming the balanced scorecard from performance measurement to strategic management. *Accounting Horizons, 15*(2), 147–160.

Kauffman, S. A. (1993). *The origins of order: Self organization and selection in evolution*. New York, NY: Oxford University Press.

Kauffman, S. A. (1996). *At home in the universe: The search for laws of complexity*. London, UK: Penguin.

Kauffman, S. A., & Levin, S. (1987). Towards a general theory of adaptive walks on rugged landscapes. *Journal of Theoretical Biology, 128*, 11–45.

Kee, R. C. (1999, December). Using economic value added with ABC to enhance production-related decision making. *Journal of Cost Management*, 3–15.

Keil, F. C. (2005). Explanation and understanding. *Annual Review of Psychology, 57*, 227–254. Retrieved from http://www.jsu.edu/dept/psychology/sebac/fac-sch/rm/Ch1-2.html

Keller, E. F., & Lloyd, E. A. (Eds.). (1992). *Keywords in evolutionary biology*. Cambridge, MA: Harvard University Press.

Kellert, S. H. (1993). *In the wake of chaos: Unpredictable order in dynamical systems*. Chicago, IL: University of Chicago Press.

Kerzner, H. (2003). *Project management: A systems approach to planning, scheduling, and controlling* (8th ed.). London, UK: Wiley.

Kim, D. (1999). *Shifting the burden: Moving beyond a reactive orientation. The systems thinker*. Waltham, MA: Pegasus Communication.

Klir, G. J. (1969). *An approach to general systems theory*. New York, NY: Van Nostrand Reinhold.

Klir, G. J. (1991). *Facets of systems science*. New York, NY: Plenum Press.

Koestler, A. (1967). *The ghost in the machine*. London, UK: Arkana.

Koestler, A. (1972). *The roots of coincidence* (2nd ed.). London, UK: Hutchinson (Paperback, 1973).

Koestler, A. (1978). *Janus; a summing up*. New York, NY: Random House.

Kotter, J. P. (1996). *Leading change*. Boston, MA: Harvard Business Press.

Koukkari, W. L., & Sothern, R. B. (2006). *Introducing biological rhythms*. New York, NY: Springer.

Kramer, D. (2010). *Management theoretician. Henri Fayol's 14 principles discussed*. http://davidkramer.wordpress.com/2010/01/21/management-thinker-henri-fayol-14-principles-discussed/

Krippendorff, K. (1986, February 2). *A dictionary of cybernetics*. Unpublished report. http://pespmc1.vub.ac.be/ASC/ALGEDO_REGUL.html

Krishnaiah, P. R., & Rao, C. R. (1988). *Quality control and reliability*. Amsterdam, Netherlands: North-Holland Ed.

Kumar, E. (2008). *Artificial intelligence*. New Delhi, India: I. K. International Publishing House.

Kurzweil, R. (2001). *The law of accelerating returns*. http://www.kurzweilai.net/the-law-of-accelerating-returns

László, E. (1983). *Systems science and world order. Selected studies*. Oxford, UK: Pergamon Press.

Latham, J. (2012). Marine cloud brightening. *Philosophical Transactions of the Royal Society*. http://rsta.royalsocietypublishing.org/content/370/1974/4217

Lee, D. (2000). The society of society: The grand finale of Niklas Luhmann. *Sociological Theory, 18*(2), 320–330.

Legion website: http://www.legion.com/legion-for-aimsun-–-vehicle-pedestrian-simulation

Leigh, J. R. (2004). *Control theory*. London, UK: The Institution of Electrical Engineers.

Lemley, B. (2002, September). The next ice age. *Discovery magazine*. http://discovermagazine.com/2002/sep/cover#.UsE5YnlTR0U

Lennox, J. G. (1992). Teleology. In E. F. Keller & E. A Lloyd (Eds.), 323–333.

Lewin, K. (1947a). Frontiers in group dynamics. In D. Cartwright (Ed.), *Field theory in social science: Selected theoretical papers* (pp. 188–237). London, UK: Tavistock.

Lewin, K. (1947b). Group decisions and social change. In T. M. Newcomb & E. L. Hartley (Eds.), *Readings in social psychology* (3rd ed., pp. 197–211). New York, NY: Holt.

Lewin, A. Y., Long, C. P., & Carroll, T. N. (1999). The coevolution of new organizational forms. *Organization Science, 10*(5), 535–550.

Lewis, M. P., Simons, G. F., & Fennig, C. D. (Eds.). (2013). *Ethnologue: Languages of the world*. Dallas, TX: SIL International. http://www.ethnologue.com/ethno_docs/distribution.asp?by = area

Li, T. Y., & Yorke, J. A. (1975). Period three implies chaos. *American Mathematical Monthly, 82*(10), 985–992.

Lienhard, J. H. (1994). I sell here, sir, what all the world desires to have. POWER. *Energy Laboratory Newsletter*, *31*, 3–9. http://www.uh.edu/engines/powersir.htm

Llinás, R., & Churchland, P. S. (1996). *The mind-brain continuum: Sensory processes*. Cambridge, MA: MIT Press.

Llinás, R., & Paré, D. (1996). The brain as a closed system modulated by senses. In R. Llinás & P. S. Churchland (Eds.), *The mind-brain continuum* (pp. 1–18). Cambridge, MA: MIT Press.

Lormand, E. (1991). *Connectionist languages of thought*. http://www-personal.umich.edu/~lormand/phil/cogsci/clot.htm

Lotka, A. J. (1925). *Elements of physical biology*. Baltimore, MD: Williams & Wilkins.

Lovelock, J. E. (1979). *GAIA, A new look at life on earth*. Oxford, UK: Oxford University Press.

Lovelock, J. E. (1988). *The ages of Gaia: A biography of our living earth*. Oxford, UK: Oxford University Press.

Lovelock, J. E. (2011). *What is Gaia*? http://www.ecolo.org/lovelock/what_is_Gaia.html

Luce, R. D., & Raiffa, H. (1967). *Games and decision*. London, UK: Wiley.

Luhman, N. (1995). *Social systems*. Stanford, CA: Stanford University Press (translation of the original book written in German language).

Luhmann, N. (1997). *Theory of society*. Stanford, CA: Stanford University Press (translation of the original book written in German language).

Mansell, W. (2011). *Control and purpose*. A Scientific and Practical Statement in Response to the UK Riots 2011. http://www.pctweb.org/LondonRiots.pdf

March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science, 2*(1), 71–87.

March, J. G., & Simon, H. A. (1958). *Organizations*. London, UK: Wiley.

Martelli, M. (1999). *Introduction to discrete dynamical systems and chaos*. London, UK: Wiley.

Maslin, M. (2004). *Global warming*. New York, NY: Oxford University Press.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review, 50*, 370–396.

Maturana, H. R., & Guiloff, G. (1980). The quest for the intelligence of intelligence. *Journal of Social and Biological Structures, 3*, 135–148.

Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition. The realization of living*. Boston, MA: Reidel Publishing (1st ed., 1972).

Maturana, H. R., & Varela, F. J. (1992). *The tree of knowledge. Biological roots of human understanding*. Boston, MA: Shambhala.

Mayer, R. E. (1992). *Thinking, problem solving, cognition* (2nd ed.). New York, NY: W. H. Freeman and Company.

McKinley, A., & Starkey, K. (Eds.). (1998). *Foucault, management and organization theory*. London, UK: Sage.

Melchizedek, D. (2004). *Dry/ice*: *Global warming revealed*. http://www.spiritofmaat.com/announce/ann_dryice.htm

Mella, P. (1997). *From systems to systems thinking: Understanding systems and thinking through systems (original: Dai sistemi al pensiero sistemico: Per capire i sistemi e pensare con i sistemi)*. Milan: Franco Angeli.

Mella, P. (2005a). Observing collectivities as simplex systems: The combinatory systems approach. *Nonlinear Dynamics, Psychology, and Life Sciences, 9*(2), 121–153.

Mella, P. (2006). Spatial co-localisation of firms and intrepreneurial dynamics: The combinatory system view. *The International Entrepreneurship and Management Journal, 2*, 391–412.

Mella, P. (2007). Combinatory systems and automata: Simulating self-organization and chaos in collective phenomena. *The International Journal of Knowledge, Culture, and Change Management, 7*(2), 17–28.

Mella, P. (2008). Observing collectivities: The combinatory systems approach in social sciences. *The International Journal of Interdisciplinary Social Sciences, 3*, 213–223.

Mella, P. (2009). *The holonic revolution. Holons, holarchies and holonic networks. The ghost in the production machine*. Pavia, Italy: University Press.

Mella, P. (2011a). Managing business value-creating organizations. *International Journal of Knowledge, Culture, and Change management, 10*, 1–18.

Mella, P. (2011b). The quality policy in value based management. *The Journal of American Academy of Business, Cambridge, 17*/2011, 68–180.

Mella, P. (2012). *Systems thinking. Intelligence in action*. New York, NY: Springer.

Mella, P. (2013). The 100 towers of medieval Pavia: A social phenomenon explained by the combinatory systems approach. *International Review of Social Sciences and Humanities, 5*(2), 168–183.

Mella, P., & Meo Colombo, C. (2012). The wheels of change in organizations. *International Journal of Knowledge, Culture, and Change Management, 11*(6), 247–265.

Mella, P., & Pellicelli, M. (2008). The origin of value based management: Five interpretative models of an unavoidable evolution. *International Journal of Knowledge, Culture, and Change management, 8*, 23–32.

Merchant, K. A. (1985). *Control in business organizations*. Boston, MA: Pitman Publishing.

Merchant, K. A., & Otley, D. T. (2007). A review of the literature on control and accountability. In C. S. Chapman, A. G. Hopwood, & M. D. Shields (Eds.), *Handbook of management accounting research* (pp. 785–804). Amsterdam, Netherlands: Elsevier Press.

Mesarovic, M., Macko, D., & Takahara, Y. (1970). *Theory of hierarchical, multi-level systems*. New York, NY: Academic Press.

Meyerhoff, J. (2005). The 20 tenets. In: *Bald ambition: A critique of Ken Wilber's theory of everything,* Ch 1(B). Frank Visser Pub. http://www.integralworld.net/meyerhoff-ba-1a.html

Miles, L. D. (1989). *Techniques of value analysis and engineering* (3rd ed.). http://wendt.library.wisc.edu/miles/book/preface.pdf

Miller, G. (2000). *Nature's fading chorus: Classic and contemporary writings on amphibians*. Washington, DC: Island Press (Google Book).

Mingers, J. (1995). *Self-producing systems: Implications and applications of autopoiesis*. New York, NY: Springer.

Mingers, J. (2002). Can social systems be autopoietic? Assessing Luhmann's social theory. *The Sociological Review, 50*(2), 278–299.

Mockler, R. J. (1970). *Readings in management control* (pp. 14–17). New York, NY: Appleton-Century-Crofts.

Modigliani, F., & Miller, M. (1958). The cost of capital, corporation finance, and the theory of investment. *American Economic Review, XLVIII*(3), 261–297.

Monod, J. (1970). *Le hazard et la nécéssité* [*Chance and necessity*]. New York, NY: Vintage Books (1972). (Original: Paris: Seuil).

Morgan, G. (1988). *Riding the waves of change: Developing managerial competencies for a turbulent world*. San Francisco, CA: Jossey-Bass.

Morin, E. (1982). *Science avec conscience*. Paris, France: Fayard. Collection "Points".

Morin, R., & Jarrel, S. (2001). *Driving shareholder value*. New York, NY: McGraw-Hill.

Morris, C. W. (1938). *Foundations of a theory of signs*. http://www.scribd.com/doc/51866596/Morris-1938-Foundations-of-Theory-of-Signs

Nagrath, I. J., & Gopal, M. (2008). *Control systems engineering*. Kent, UK: Anshan.

NASA. (2004). A chilling possibility. *Science*. http://science1.nasa.gov/science-news/science-at-nasa/2004/05mar_arctic/

Neuman, Y. (2003). *Processes and boundaries of the mind*. New York, NY: Springer.

Newell, A., & Simon, H. A. (1976). Computer science as empirical enquiry: Symbols and search. *Communications of the Association for Computing Machinery, 19*(3), 113–126.

Nickols, F. (2010). *Change management 101. A primer*. http://www.nickols.us/change.pdf

Nickols, F. (2011). Reengineering the problem-solving process. Finding better solutions faster. Retrieved from, www.nickols.us/reengineering.pdf

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science, 5*(1), 14–37.

Nöth, W. (1995). *Handbook of semiotics*. Bloomington, IN: Indiana University Press.

Nystrom, P., & Starbuck, W. (Eds.). (1981). *Handbook of organizational design*. Oxford, UK: Oxford University Press.

Ogata, K. (2002). *Modern control engineering*. Englewood Cliffs, NJ: Prentice Hall.

Olve, N. G., Roy, J., & Wetter, M. (1999). *Performance drivers: A practical guide to using the balanced scorecard*. London, NY: Wiley.

Osborn, A. F. A. (1953). *Applied imagination: The principles and procedures of creative thinking*. New York, NY: Scribner's sons.

Ostrom, E. (1990). *Governing the commons. The evolution of institutions for collective action*. Cambridge, UK: Cambridge University Press.

Otley, D. (1999). Performance management: A framework for management control systems research. *Management Accounting Research, 10*, 363–382.

Ouchi, W. G., & Maguire, M. A. (1975). Organizational control. Two functions. *Administrative Science Quarterly, 20*, 559–569.

Parsons, T. (1951). *The social system*. New York, NY: Free Press.

Parsons, T. (1971). *The system of modern societies*. Englewood Cliffs, NJ: Prentice Hall.

Perkel, D. H. (1993). Computational neuroscience. Scope and structure. In E. L. Schwartz (Ed.), *Computational neuroscience* (pp. 38–45). Cambridge, MA: MIT Press.

Pichler, F. (2000). On the construction of A. Koestler's holarchical networks. In R. Trappl (Ed.), *Cybernetics and systems* (1st ed., pp. 80–84).

Plamondon, A. L. (1979). *Whitehead's organic philosophy of science*. New York, NY: State University of New York Press.

PlanIT Valley project. (2013). http://www.living-planit.com/design_wins.htm

Polanyi, M. (1967). *The tacit dimension*. Chicago, IL: University of Chicago Press.

Porter, M. E. (2008). *On competition (updated and expanded edition)*. Boston, MA: Harvard Business School Press.

Powers, W. T. (1973). *Behavior: The control of perception*. Chicago, IL: Aldine.

Prahalad, C. K., & Bettis, R. A. (1986). The dominant logic: The new linkage between firm diversity and performance. *Strategic Management Journal, 7*(6), 485–501.

Prieto, L. (1966). *Messages et signaux [Messages and signals]*. Paris, France: P.U.F.

Proctor, T. (1999). *Creative problem solving for managers*. New York, NY: Taylor & Francis.

Rao, A., & Georgeff, M. (1992). Social plans: preliminary report. In E. Werner & C. Castelfranchi (Eds.), *Decentralized AI 3, Proceedings of MAAMAW'91*. Amsterdam, Netherlands: Elsevier North Holland.

Richardson, L. F. (1949). *Arms and insecurity*. London, UK: Stevens & Sons (published posthumously, 1960).

Richardson, G. P. (1991). *Feedback thought in social science and system theory*. Philadelphia, PA: University of Pennsylvania Press.

Richmond, B. (1991). *Systems thinking. Four key questions*. Watkinsville, GA: High Performance Systems.

Richmond, B. (1993). Systems thinking: critical thinking skills for the 1990s and beyond. *System Dynamics Review, 9*(2), 113–133.

Richmond, B. (1994). System dynamics/systems thinking: Let's just get on with it. *International Systems Dynamics Conference, Sterling, Scotland*. Retrieved from http://webspace.webring.com/people/ah/himadri_banerji/pdf/systhnk.pdf

Riedel, S. (2005). Edward Jenner and the history of smallpox and vaccination. *Proceedings of Baylor University Medical Center, 18*(1), 21–25. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1200696/

Roberts, N. (1978). Teaching dynamic feedback systems thinking: An elementary view. *Management Science, 24*(8), 836–843.

Roberts, N., Andersen, D., Garet, M. S., Shaffer, W. A., & Deal, R. M. (1983). *Introduction to computer simulation: A system dynamics modeling approach*. Reading, MA: Addison Wesley.

Robertson, S. I. (2001). *Problem solving*. Hove, East Sussex: Psychology Press.

Rosheim, M. (2006). *Leonardo's lost robots*. New York, NY: Springer.

Rumelhart, D., & McClelland, J. (1982). An interactive activation model of context effects in letter perception. *Psychological Review*, *89*, 60–94. http://psych.stanford.edu/~jlm/papers/RumelhartMcClelland82.pdf.

Russell, S., & Norvig, P. (2009). *Artificial intelligence, a modern approach* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Salmon, W. C. (1990). *Four decades of scientific explanation*. Minneapolis, MN: University of Minnesota Press.

Sandquist, G. M. (1985). *Introduction to system science*. Englewood Cliffs, NJ: Prentice Hall.

Saussure, F. de. (1916). Course in general linguistics. (French: Cours de linguistique générale). Charles Bally, & Albert Sechehaye (Eds.). New York, NY: Philosophical Library, 1959. Retrieved from, http://faculty.smu.edu/dfoster/cf3324/saussure.htm

Scott, S. G., & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *The Academy of Management Journal, 37*(3), 580–607.

Segre, S. (2012). *Talcott parsons: An introduction*. Lanham, MD: University Press of America.

Senge, P. (1990). *The fifth discipline: The Art and practice of the learning organization* (1st ed.). New York, NY: Doubleday/Currency.

Senge, P. (1997). *Foreword* to A. De Geus, (1997).

Senge, P. (2006). *The fifth discipline: The Art and practice of the learning organization*. New York, NY: Doubleday/Currency (last edition, revised and enlarged).

Senge, P., Kleiner, A., Roberts, C., Ross, R., & Smith, B. (1994). *The fifth discipline fieldbook*. New York, NY: Doubleday.

Senge, P., & Lannon-Kim, C. (1991) The systems thinking approach. *The Systems Thinker Newsletter*, *2*(5). Cambridge MA: Kendall Square.

Senge, P., Smith, B., Kruschwitz, N., Laur, J., & Schley, S. (2008). *The necessary revolution. Working together to create a sustainable world*. New York, NY: Broadway Books.

SFN, Society for neuroscience. http://www.sfn.org/about/about-neuroscience

Shaikh, S., & Leonard-Amodeo, J. (2005). The deviating eyes of Michelangelo's David. *Journal of the Royal Society of Medicine*, *98*(2), 75–76. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1079389/

Shibata, H. (1998). *Problem solving: Definition, terminology, and patterns*. www.mediafrontier.com/Article/PS/PS.htm

SHRM, Society for Human Resource Management. http://www.shrm.org/TemplatesTools/Glossaries/HRTerms/Pages/c.aspx

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society, 106*(6), 467–482.

Simon, H. A., & Associates. (1986). *Decision making and problem solving*. Washington, DC: National Academy of Sciences, National Academy Press. Available online at http://dieoff.org/page163.htm

Simons, R. (1995). *Levers of control: How managers use innovative control systems to drive strategic renewal*. Boston, MA: Harvard Business School Press.

Skyttner, L. (2005). *General systems theory: Problems, perspectives, practice*. Singapore: World Scientific Pub. Online Google Books.

Smith, B. S. (2013). Zeno's paradox of the tortoise and achilles. *PRIME*. http://platonicrealms.com/encyclopedia/Zenos-Paradox-of-the-Tortoise-and-Achilles/

Söderbaum, P. (2008). *Understanding sustainability economics: Towards pluralism in economics*. London, UK: Earthscan.

Sontag, E. D. (1998). *Mathematical control theory: Deterministic finite dimensional systems* (2nd ed.). New York, NY: Springer.

Spelta, A. M. (1602). History written by Antonio Maria Spelta, citizen of Pavia (original: Historia d'Antonio Maria Spelta, Cittadino Pavese), Pavia: Pietro Bartoli, MDCII (Library of the University of Pavia).

Sterman, J. D. (2000). *Business dynamics: Systems thinking and modeling for a complex world*. New York, NY: McGraw-Hill.

Sterman, J. D. (2001). System dynamics modeling: Tools for learning in a complex world. *California Management Review, 43*(4), 8–25.

Stern, J. M., Bennett Stewart, G., III, & Chew, D., Jr. (1995). The EVA financial system. *Journal of Applied Corporate Finance, 3*(2), 38–55.

Takeuchi, Y. (1995). *Global dynamical properties of Lotka-Volterra systems*. Singapore: World Scientific Pub.

Tannenbaum, A. S. (1962). Control in organizations: individual adjustment and organizational performance. *Administrative Science Quarterly, 7*(2), 236–257.

Tarski, A. (1944). The semantic conception of truth and the foundations of semantics. *Philosophy and Phenomenological Research, 4*(3), 341–376. http://www.ditext.com/tarski/tarski.html

The Dalai Lama (1998). The Four Noble Truths, Thorsons (Ed). (Paperback - Jan 1998).

The Dalai Lama (2010). Four Noble Truths. Retrieved from www.lamayeshe.com/index.php?sect=article&id=380

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59, 433–460. http://loebner.net/Prizef/TuringArticle.html

Uribe, R. (1981). Modeling autopoiesis. In M. Zeleny (Ed.), *Autopoiesis: A theory of living organization*. New York, NY: Elsevier-North Holland.

Urban Dharma. Buddhism in America. Retrieved from www.urbandharma.org/pdf/PDF_BuddhismCourse/03_%20Four%20Noble%20Truths.pdf

Vanecek, A., & Celikovsky, S. (1996). *Control systems: From linear analysis to synthesis of chaos*. Englewood Cliffs, NJ: Prentice Hall.

Varela, F. J. (1979). *Principles of biological autonomy*. Boston, MA: Kluwer Academic.

Varela, F. J. (1981). Describing the logic of living. In M. Zeleny (Ed.), *Autopoiesis: A theory of living organization* (pp. 36–48). New York, NY: Elsevier.

Varmah, K. R. (2010). *Control systems*. New Delhi, India: Tata McGraw-Hill Education.

Velentzas, J., & Broni, G. (2011). Cybernetics and autopoiesis theory as a study of complex organizations. A systemic approach. *ICOAE, International Conference on Applied Economics* (pp. 737–747).

Vicari, S. (1991). *L'impresa vivente [The firm as a living system]*. Milan, Italy: Etas Ed.

Visilab. (2013). *Blind assistant project*. http://visilab.unime.it/~filippo/BlindAssistant/BlindAssistant.htm

Volterra, V. (1926). Variations and fluctuations of the number of individuals in animal species living together. Animal Ecology. New York, NY: McGraw-Hill (1931) (Original: Variazioni e fluttuazioni del numero d'individui in specie animali conviventi. *Atti dell'Accademia Nazionale* dei *Lincei*, 31–113).

Volterra, V. (1931). *Leçons sur la théorie mathématique de la Lutte pour la vie [Lessons on the mathematical theory of struggle for life]*. Paris, France: Gauthier-Villars.

von Bertalanffy, L. (1968). *General system theory: Foundations, development, applications*. New York, NY: Braziller.

von der Pütten, J. C. (2008). *Moral issues and concerns about China's one-child policy: A cosmopolitan perspective*. Norderstedt, Germany: GRIN.

von Foerster, H. (1984). *On constructing a reality*. In P. Watzlawick (Ed., 1984). http://ada.evergreen.edu/~arunc/texts/InventingSystems/readings2.pdf

von Foerster, H. (1990). Ethics and second-order cybernetics. *Cybernetics & Human Knowing*. http://www.imprint.co.uk/C&HK/vol1/v1-1hvf.htm

von Foerster, H. (2003). *Understanding understanding: essays on cybernetics and cognition*. New York, NY: Springer.

von Krogh, G., & Roos, J. (1996). The epistemological challenge: Managing knowledge and intellectual capital. *European Management Journal, 14*(4), 333–337.

von Neuman, J., & Morgenstern, O. (1953). *Theory of games and economic behavior* (5th ed.). Princeton, NJ: Princeton University Press (1st ed., 1944).

Waldrop, M. M. (1992). *Complexity. The emerging science at the edge of order and chaos*. New York, NY: Simon & Shuster.

Walras, L. (1874). *Elements of pure economics: Or the theory of social wealth (translated in French in 1926 and then published in English)*. Homewood, IL: Irwin. 1954.

Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization Science, 6*(3), 280–321.

Wasson, C. S. (2006). *System analysis, design, and development*. London, UK: Wiley.

Waterman, R., Jr., Peters, T., & Phillips, J. R. (1980). Structure is not organisation. *Business Horizons, 23*(3), 14–26.

Waters, C. D. (1992). *Inventory control and management*. London, UK: Wiley.

Watson, A. J., & Lovelock, J. E. (1983). Biological homeostasis of the global environment: The parable of Daisyworld. *Tellus B, 35*(9), 286–289.

Watzlawick, P. (1984). *The invented reality: How do we know what we believe we know? (Contributions to constructivism)*. New York, NY: Norton & Co Inc.

Watzlawick, P., Weakland, J. H., & Fisch, R. (1974). *Change: Principles of problem formation and problem resolution*. New York, NY: Norton.

Wellemin, J. H. (1990). *Customer satisfaction through quality*. Brickley, UK: Chartwell-Bratt (Publishing & Training) Ltd.

Widrow, B., Gupta, N. K., & Maitra, S. (1973). Punish/reward: Learning with a critic in adaptive threshold systems. *IEEE Transactions on Systems, Man, and Cybernetics, 5*, 455–465.

Wiener, N. (1954). *The human use of human beings: cybernetics and society*. Boston, MA: Houghton Mifflin (Original work published in 1949).

Wiener, N. (1961). *Cybernetics: Or control and communication in the animal and the machine*. Cambridge, MA: MIT Press. Google Books. (Original work published 1948).

Wikipedia. (2013). *List of wars 2011/present*. http://en.wikipedia.org/wiki/List_of_wars_2011–present

Wilber, K. (2000). *Sex, ecology, spirituality: The spirit of evolution*. Boston, MA: Shambhala (1st ed., 1995).

Wilber, K. (2001). *A brief history of everything*. Boston, Ma: Shambhala (1st ed., 1996).

Wilson, R. H. (1934). *A scientific routine for stock control*. Bell telephone system technical publications: Monographs

Wooldridge, M., & Jennings, N. R. (1994). Toward a theory of cooperative problem solving. In *Proceedings of the 6th European workshop on modelling autonomous agents in a multi-agent world*, Odense, Denmark, August.

Woolf, M. B. (2007). *Faster construction projects with CPM scheduling*. New York, NY: McGraw-Hill.

Woolsey, R., & Maurer, R. (2005). *Inventory control*. Marietta, GA: Lionheart Publishing.

Young, S. D. (1999). Some reflections on accounting adjustments and economic value-added. *Journal of Financial Statement Analysis, 4*(2), 7–19.

Zeleny, M. (1981). *Autopoiesis, a theory of living organization*. New York, NY: Elsevier-North Holland.

Zeleny, M., & Hufford, K. D. (1992). The application of autopoiesis in systems analysis: Are autopoietic systems also social systems? *International Journal of General Systems, 21*, 145–160.

Zhao, W., & Chellappa, R. (2006). *Face processing. Advanced modelling and methods*. Boston, MA: Academic Press.

Zhong, J. (2006). *PID controller tuning: A short tutorial*. https://www.yumpu.com/en/document/view/4531382/pid-controller-tuning-a-short-tutorial

Zuradelli, C. (1888). The towers of Pavia (original: Le torri di Pavia), Pavia: Tipografia F.lli Fusi (Library of the University of Pavia).

Index